

Release 3.1A John F. Collins, Biocomputing Research Unit
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

```
Title: >US-09-316-163-9
Description: (1-207) from US09316163.pep
Perfect Score: 1573 ✓
Sequence: 1 EDCNELPPRRNTEILTGSM.....
```

Searched: 142080 seqs, 47172406 residues

Database: pir62

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	1573	100.0	449	1	NBHUHS	complement factor H p	0.00e+00
2	1573	100.0	1231	1	NBHUH	complement factor H p	0.00e+00
3	1154	73.4	1234	1	NBMKS	complement factor H p	1.13e-20
4	679	43.2	669	2	S65551	factor H - bovine (fr	5.85e-1
5	503	32.0	1053	2	S46199	probable complement r	6.02e-99
6	334	21.2	597	1	S53711	C4B alpha chain prec	6.02e-55
7	310	19.7	560	2	T16833	hyolectin - protein	3.60e-44
8	308	19.6	482	2	TC5092	F-selectin - pig	1.08e-4
9	308	19.6	830	2	A30359	P-selectin precursor	1.08e-4
10	303	19.3	289	1	WMVZSP	apolipoprotein H homo	1.69e-44
11	303	19.3	768	2	A42755	P-selectin H precursor	1.69e-44
12	303	19.3	768	2	IF5831	P-selectin - rat	1.69e-44
13	297	18.9	345	1	NBM81	apolipoprotein H prec	4.55e-44
14	298	18.9	610	2	A35046	endothelial leukocyte	2.63e-44
15	288	18.9	610	2	A32606	endothelial leukocyte	2.63e-44
16	282	18.6	345	1	NBBO	apolipoprotein H prec	7.01e-44
17	283	18.6	381	1	A26339	decay-accelerating fa	4.06e-44
18	293	18.6	440	2	A26359	decay-accelerating fa	4.06e-44
19	290	18.4	551	2	I46709	endothelial leukocyte	2.09e-44
20	290	18.4	597	1	NBHUC4	C4b-binding protein a	2.09e-44
21	288	18.3	302	1	WMBE2E	secretory complement	6.22e-44
22	288	18.3	360	1	WMBE2E	membrane bound comple	6.22e-44
23	288	18.3	612	2	S23174	endothelial leukocyte	6.22e-44

25	287	18.2	55.1	2	I46708	endothelial leukocyte
24	284	18.1	340	2	I56234	decay-accelerating fa
26	282	17.9	263	1	C38838	complement control pr
27	281	17.9	349	2	G04921	sperm CD46 - human (f
28	281	17.9	359	2	I57998	membrane cofactor pro
29	281	17.9	377	2	I54479	membrane cofactor pro
30	281	17.9	301896	3	S01896	membrane cofactor pro
31	280	17.8	B47255	6	B47255	E-selectin precursor
32	279	17.7	618	2	A45900	complement C3b recept
33	276	17.5	473	2	B38738	coagulation factor C-
34	276	17.5	1019	2	A38738	coagulation factor C
35	272	17.3	352	2	C5J194	membrane cofactor pro
36	272	17.3	359	2	C5J138	membrane cofactor pro
37	270	17.2	345	1	NBHJ	apolipoprotein H prec
38	270	17.2	610	1	I46001	C4b-binding protein a
39	270	17.2	1091	1	PL0009	complement C3d/Epstei
40	267	17.0	2039	1	A28307	complement C3b/C4d re
41	267	17.0	2489	2	I37012	complement C3b/C4d re
42	266	16.9	4485	2	S67702	E-selectin - bovine
43	266	16.9	4477	2	C42054	complement regulatory
44	264	16.8	345	1	JN0465	apolipoprotein H prec
45	265	16.8	469	1	NBMSCA	C4b-binding protein a

ALIGNMENTS

```

RESULT      1
ENTRY
TITLE
ATTN       NBH0HS      #type complete
TITLE      complement factor H precursor, short splice form - human
ORGANISM   complement factor H-related protein: complement protein H
DATE       31-Dec-1993 #sequence_revision 23-Feb-1996 #text_change
          22-Jun-1999
REFERENCES
#authors   S03013; B60238; A27877; A61103; A26505; S10679
#journal   S00254
#title     Ripoche, J.; Day, A.-C.; Harris, T.J.B.; Slim, R.B.
           Biochem. J. (1988) 249:593-602
           The complete amino acid sequence of human complement factor
           H.
#cross-references MUID:88134059
#accession S03013
#molecule_type mRNA
#residues 1-449 ##label RIP
##cross-references EMBL:X07523; EMBL:Y00716; NID:g32492;
           PIDN:CA30403.1; PID:g758073
           part of this sequence, including the amino end of the
           mature protein was confirmed by protein sequencing
           402-Tyr was also found
#note
A60238
#authors   Estaller, C.; Schwaebler, W.; Dierich, M.; Weiss, E.H.
#journal   Eur. J. Immunol. (1991) 21:799-802
#title     Human complement factor H: two factor H proteins are derived
           from alternatively spliced transcripts.
#cross-references MUID:91184292
#accession B60238
#status    not compared with conceptual translation
#molecule_type mRNA
#residues 1-33,434-449 ##label EST
#note      Only portions of this 1.8 kilobase mRNA were sequenced
A27877
#authors   Schultz, T.F.; Schwaebler, W.; Stanley, K.K.; Weiss, E.;
           Dierich, M.P.
#journal   Eur. J. Immunol. (1986) 16:1351-1355
#title     Human complement factor H: Isolation of cDNA clones and
           partial cDNA sequence of the 38-kDa tryptic fragment
           containing the binding site for C3b.
#cross-references MUID:87054207
#accession A27877
#molecule_type mRNA
#residues 'IL',55-401,'Y',403-449 ##label SCH
#note      an additional nucleotide present within the codon for
           Glu-310 was thought to be a cloning artifact and was
           ignored in translation

```

```

#AUTHORS
#authors Schwäbke, W.; Zwirner, J.; Schulz, T.F.; Linke, R.P.;
#journal Eur. J. Immunol. (1987) 17:1485-1489
#title Human complement factor H: expression of an additional
#cross-references MUID:88055295 truncated gene product of 43 kDa in human liver.
#accession A61103
##status not compared with conceptual translation
##molecule_type mRNA
##residues 27-76 ##label SC2
##note this is a partial sequence of an alternatively spliced
1.8 kilobase mRNA that is translated to yield a 43 k
form related to factor H

REFERENCE
#authors A26505
#journal Sim, R.B.; Discipio, R.G.
#title Biochem. J. (1982) 205:285-293
#cross-references MUID:83048213 Purification and structural studies on the complement-system
#accession A26505 control protein beta-1-H (factor H).

REFERENCE
#molecule_type protein
#residues 19-20,'Q',22-29,'V',31-33,'Q',35 ##label SIM
#journal A44551
#authors Barlow, P.N.; Norman, D.G.; Steinkasserer, A.; Horne, T.J.;
#title Pearce, J.; Driscoll, P.C.; Sim, R.B.; Campbell, I.D.
#cross-references MUID:92232649 Biochemistry (1992) 31:3626-3634
#contents Solution structure of the fifth repeat of factor H: A second
#accession MUID:92232649 example of the complement control protein module.
#authors Kristensen, T.; Wetzel, R.A.; Tack, B.F.
#journal J. Immunol. (1986) 136:3407-3411
#title Structural analysis of human complement protein H: homology
#cross-references MUID:86169701 with C4b binding protein, beta(2)-glycoprotein I, and the
#accession S10479 Ba fragment of B.

#molecule_type mRNA
#residues 226-408,'Y',403-449 ##label KRI
#cross-references GB:M12383; NID:G180472; PID:AAA52013.1; PID:G180473
#comment Factor H has also been found bound to cell membranes in an unknown
#cross-references MUID:86169701 manner. However, it has at least one cell attachment site motif
#comment in repeat 4.
#comment Alternative transcripts of 4.3, 1.8, and 1.4 kilobases are expressed
#comment in liver. See also PIR:MB0UH.

GENETICS
#gene HFI
#cross-references GDB:120041; OMIM:134370 GDB:HFI; HF
#map_position 1q32-1q32
GENETICS
#gene GDB:HF2; HF
#cross-references GDB:129095
#map_position 1q32-1q32
#note the correspondence between the two loci and the sequences
indicated is unclear; factor H has been reported to have
several allelic forms

FUNCTION
#description a cofactor in the inactivation of C3b by serine proteinase I;
also increases the rate of dissociation of the C3bBb
complex (C3 convertase) and the (C3b)NB complex (C5
convertase) in the alternative complement pathway
#pathway complement alternate pathway
#classification #superfamily complement factor H; complement factor H repeat
homology
#keywords alternative splicing; complement alternate pathway;
glycoprotein; plasma

FEATURE
1-18 #domain signal sequence #status predicted #label SIG\
19-449 #product complement factor H, short splice form #status
experimental #label MNT\
21-80 #domain complement factor H repeat homology #label FH01\

```

```

85-141      #domain complement factor H repeat homology #label FH02
146-205     #domain complement factor H repeat homology #label FH03
210-282     #domain complement factor H repeat homology #label FH04
246-248     #region cell attachment (R-G-D) motif\
267-320     #domain complement factor H repeat homology #label FH05
323-385     #domain complement factor H repeat homology #label FH06
389-442     #domain complement factor H repeat homology #label FH07
21-66,52-80,85-129,
114-141,146-192,
178-205,210-251,
237-282,267-309,
294-320,325-374,
357-385,389-431,
416-442
#disulfide_bonds #status predicted\
#binding_site carbohydrate (Asn) (covalent) #status absent
SUMMARY      #length 449 #molecular_weight 51007 #checksum 6077

Query Match      100.0%; Score 1573; DB 1; Length 449;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db    19 EDCNCLPPRRMTTEILTGSMSQTYPESTGALHYKCRPRYSLGWIVWCRCGEWVALNPRLR 78
      |||
      |EDCNCLPPRRMTTEILTGSMSDQITPREGIQAITYKCRPRYSLGNVWCRCGEWVALNPRLR 60
Qy    1 KQKRPCGHDPDFPGFTTLTGWNVEFYGVKAYTCNEGYTOLGLGEINYEACDTDGWTNDI 138
      |||
      |KQKRPCGHDPDFPGFTTLTGWNVEFYGVKAYTCNEGYTOLGLGEINYEACDTDGWTNDI 120
Db    139 PICEVVKKLPYTAENCKIVSSAEPREHFEQAVRFVCNSGRTKEGDENHCSDDGFW 198
      |||
      |PICEVVKKLPYTAENCKIVSSAEPREHFEQAVRFVCNSGRTKEGDENHCSDDGFW 180
Qy    121 SKERPKCWEISCKSPDYINGSPISOKI 225
      |||
      |SKERPKCWEISCKSPDYINGSPISOKI 207
Db    199 SKERPKCWEISCKSPDYINGSPISOKI 225
      |||
      |SKERPKCWEISCKSPDYINGSPISOKI 207
Qy    181 SKERPKCWEISCKSPDYINGSPISOKI 207
      |||
      |SKERPKCWEISCKSPDYINGSPISOKI 207

RESULT        2
ENTRY         NBHDH                      #type complete
TITLE         Complement factor H precursor, long splice form - human
ORGANISM      Homo sapiens #common_name man
DATE          31-Dec-1993 #sequence_revision 31-Dec-1993 #ext_change
              22-Jun-1999
ACCESSIONS   S00254; A60726; A61555; A26505; I72654; S66298
REFERENCE     S00254
#authors      Ripoche, J.; Day, A.J.; Harris, T.D.R.; Sim, R.B.
#journal      Blochem. J. (1988) 249:593-602
#title        The complete amino acid sequence of human complement factor H.
#cross-references NUID:88134059
#accession    S00254
##molecule_type mRNA
##residues    1-1231 ##label RIP
##cross-references EMBL:Y00716; NID:g31964; PIDN:CMA68704.1; PID:g31965
#note         402-Tyr was also found
#note         parts of this sequence, including the amino and carboxy
#note         ends of the mature protein, were confirmed by protein
#note         sequencing
REFERENCE     A60238
#authors      Estaller, C.; Schwaebler, W.; Dietrich, M.; Weiss, E.H.
#journal      Eur. J. Immunol. (1991) 21:799-802
#title        Human complement factor H: two factor H proteins are derived
#title        from alternatively spliced transcripts.
#cross-references NUID:911844292
#accession    A60238
#status       not compared with conceptual translation
##molecule_type mRNA
##residues    1-56;1177-1231 ##label EST
##note         only portions of this 4.3 kilobase mRNA were sequenced
REFERENCE     A54726
#authors      Day, A.J.; Ripoche, J.; Lyons, A.; McIntosh, B.; Harris,
```

#journal T.J.R.: Sim, R.B. Biosci. Rep. (1987) 7:201-207
 #title Sequence analysis of a cDNA clone encoding the C-terminal end of human complement factor H.
 #cross-references MUID:88025472
 #accession A54726
 #status not compared with conceptual translation
 #molecule-type mRNA
 #residues 'DEFN', 579-1231 #label DAY
 #cross-references GB:M17517; NID:g180497; PIDN:AAA52016.1; PID:g180498
 #note parts of this sequence were determined by protein sequencing
 REFERENCE
 #authors A61565
 #journal Ripocher, J.; Day, A.J.; Willis, A.C.; Belt, K.T.; Campbell, R.D.; Sim, R.B.
 #title Biosci. Rep. (1986) 6:65-72
 #cross-references MUID:86188123
 #accession A61565
 #status not compared with conceptual translation
 #molecule-type mRNA
 #residues 'METGRNHLNAKI', 1050-1057, 'T', 1059-1102 #label R12
 REFERENCE
 #authors A26505
 #journal Sim, R.B.; Discipio, R.G.
 #title Biochem. J. (1982) 205:285-293
 #cross-references MUID:83048213
 #accession A26505
 #molecule-type protein
 #residues 19-20, 'Q', 22-29, 'V', 31-33, 'Q', 35 #label SIM
 REFERENCE
 #authors A4551
 #journal Barlow, P.N.; Norman, D.G.; Steinkasserer, A.; Horne, T.J.; Pearce, J.; Driscoll, P.C.; Sim, R.B.; Campbell, I.D.
 #title Biochemistry (1992) 31:3626-3634
 #cross-references MUID:92232649
 #accession A49224
 #status annotation: NMR structure determination, residues 264-292
 #journal Norman, D.G.; Barlow, P.N.; Baron, M.; Day, A.J.; Sim, R.B.; Campbell, I.D.
 #title J. Mol. Biol. (1991) 219:717-725
 #cross-references MUID:91278097
 #accession I56100
 #status annotation: NMR structure determination, residues 927-985
 #journal Esteller, C.; Koistinen, V.; Schwaible, W.; Dierich, M.P.; Weiss, E.H.
 #title J. Immunol. (1991) 146:3190-3196
 #cross-references MUID:91278097
 #accession I72654
 #status translated from GB/EMBL/DBJ
 #molecule-type mRNA
 #residues 1047-1231 #label RES
 #cross-references GB:M55294; NID:g183766; PIDN:AAA35948.1; PID:g183767
 REFERENCE
 #authors S66298
 #journal Caron, J.A.; Bates, R.C.; Smith, A.I.; Tetz, T.; Arellano, A.; Gordon, D.L.; Burns, G.F.
 #title Biochim. Biophys. Acta (1996) 1289:305-311
 #cross-references MUID:96205365
 #accession S66298
 #status preliminary
 #molecule-type protein

#residues 411-419:574-578,580-582 #label CAR
 COMMENT Factor H has also been found bound to cell membranes in an unknown manner. However, it has at least one cell attachment site motif in repeat 4.
 COMMENT Alternative transcripts of 4.3, 1.8, and 1.4 kilobases are expressed in liver. See also PIR:NBH0HS.
 GENETICS
 #gene HFI
 #cross-references GDB:120041; OMIM:134370
 #map-position 1q32-1q32
 GENETICS
 #gene HFI
 #cross-references GDB:129095
 #map-position 1q32-1q32
 #note the correspondence between the two loci and the sequences indicated is unclear; factor H has been reported to have several allelic forms
 FUNCTION
 #description a cofactor in the inactivation of C3b by serine proteinase I; also increases the rate of dissociation of the C3bB complex (C3 convertase) and the (C3b)nb complex (C5 convertase) in the alternative complement pathway
 CLASSIFICATION
 #pathway complement alternate pathway
 #superfamily complement factor H; complement factor H repeat homology
 #alternative splicing: complement alternate pathway; glycoprotein; plasma
 KEYWORDS
 #domain signal sequence #status predicted #label SIG
 #product complement factor H #status experimental #label MP1
 #product complement factor H, short splice form #status experimental #label MAR
 #domain complement factor H repeat homology #label FH01
 #domain complement factor H repeat homology #label FH02
 #domain complement factor H repeat homology #label FH03
 #domain complement factor H repeat homology #label FH04
 #domain complement factor H repeat homology #label FH05
 #domain complement factor H repeat homology #label FH06
 #domain complement factor H repeat homology #label FH07
 #domain complement factor H repeat homology #label FH08
 #domain complement factor H repeat homology #label FH09
 #domain complement factor H repeat homology #label FH10
 #domain complement factor H repeat homology #label FH11
 #domain complement factor H repeat homology #label FH12
 #domain complement factor H repeat homology #label FH13
 #domain complement factor H repeat homology #label FH14
 #domain complement factor H repeat homology #label FH15
 #domain complement factor H repeat homology #label FH16
 #domain complement factor H repeat homology #label FH17
 #domain complement factor H repeat homology #label FH18
 #domain complement factor H repeat homology #label FH19
 #domain complement factor H repeat homology #label FH20


```

DATE                31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change
ACCESSIONS         JC5092
REFERENCE           Winbler, H.; Brostjan, C.; Szilmadia, V.; Nataraajan, G.;
#authors            Anrather, J.; Bach, F.H.
#journal             Gene (1996) 176:67-72
#title              The intron-exon structure of the porcine E-selectin-encoding
                    gene.
GENETICS            #cross-references MIMID:97075911
                    #contents endothelial cells
                    #accession JC5092
                    ##molecule-type DNA
                    ##residues 1-482 #label WIN
COMMENT              #cross-references GB:U37521; NID:g1052974; PID:g1052975
                    This protein is a member of the selectin family of adhesion
                    molecules.
CLASSIFICATION      13/1; 142/1; 178/1; 237/1; 300/1; 363/1; 422/1; 459/1; 466/1
FEATURE             #superfamily complement factor H repeat homology; C-type
                    lectin homology
13-139              #domain C-type lectin homology #label LCH\
181-235              #domain complement factor H repeat homology #label FH1\
240-298              #domain complement factor H repeat homology #label FH2\
303-361              #domain complement factor H repeat homology #label FH3\
366-420              #domain complement factor H repeat homology #label FH4\
SUMMARY             #length 482 #molecular_weight 53341 #checksum 7729
Query Match          19.6%; Score 308; DB 2; Length 482;
Best Local Similarity 28.9%; Pred. No. 1,08e-47;
Matches 52; Conservative 38; Mismatches 82; Indels 8; Gaps 6;
Db                   196 OSLPNTTCAFCCKGFEELIHOCTSSGSDPKRP--TCNAVCDTVGHPRQNDVSC 253
OY                   22 QYTBETAQAIYICRGYKSLGNVTMVC-RKGWVALNPLRKCRCRCPHGPTPTFTFL 80
                        :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db                   254 NHSSGEFAYKSTCHFTCAEGFGLOGPAQT-ECTAOGQWTQAQPVCAKPCAVASOPKNK 312
OY                   81 TGGNVEFYGVXAV-YTCNEGYYQLLGEINREDDTDG-WTNIDPICEYVKCLPYAPENG 137
                        :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db                   313 -LVKRTHSPTGFNYKKSCAFSGCEGELNGSQLACTSGQGOMTOEVPSCQVVCCSLEY 371
OY                   138 KIVSSAMPEDEHYHGGQVRFVCSGYKIIEGDMEHCSDDGFWMSKEKPVCIVEISCKSPDV 197
RESULT              9
ENTRY               A30359 #type complete
TITLE               P-selectin precursor-human
ALTERNATE_NAMES     CD62 antigen; granule membrane protein 140 *
ORGANISM            Homo sapiens #common_name man
DATE                18-Oct-1989 #sequence_revision 30-Sep-1991 #text_change
                    29-Aug-1997
REFERENCES           A30359
                    A30359 Johnston, G.I.; Cook, R.G.; McEver, R.P.
                    #authors Cell (1989) 56:1033-1044
                    #journal Cloning of GMP-140, a granule membrane protein of platelets
                    #title and endothelium: sequence similarity to proteins involved
                    in cell adhesion and inflammation.
                    #cross-references MIMID:89168432
                    #accession A30359
GENETICS            ##molecule-type mRNA
                    ##residues 1-830 #label JOH
                    ##cross-references GB:M25322
                    ##note parts of this sequence, including the amino end of the
                    mature protein, were confirmed by protein sequencing
CLASSIFICATION      GDB:SELP; GRMP
                    #gene
                    #map_position 1q22-1q25
                    #map_cross-references GDB:1120018; OMIM:173610
CLASSIFICATION      #superfamily complement factor H repeat homology; EGF
                    homology

```

KEYWORDS

cell adhesion; glycoprotein; phosphatidyl; phosphoprotein; surface antigen; transmembrane protein

FEATURE

1-41 #domain signal sequence #status predicted #label SIG
42-830 #product P-selectin #status experimental #label MAT
163-194 #domain complement factor H repeat homology #label FH01
200-257 #domain complement factor H repeat homology #label FH02
262-319 #domain complement factor H repeat homology #label FH03
324-381 #domain complement factor H repeat homology #label FH04
386-443 #domain complement factor H repeat homology #label FH05
448-505 #domain complement factor H repeat homology #label FH06
510-567 #domain complement factor H repeat homology #label FH07
572-629 #domain complement factor H repeat homology #label FH08
642-699 #domain complement factor H repeat homology #label FH09
704-761 #domain complement factor H repeat homology #label TMN
772-795 #domain transmembrane #status predicted #label TMN
796-830 #domain intracellular #status predicted #label CYT
54,98,180,212,219,
411,460,518,665,
716,723,741

SUMMARY

#length 830 #molecular-weight 90754 #checksum 2794
Query Match 19.6%; Score 308; DB 2; Length 830;
Best Local Similarity 26.8%; Pred. No. 1.08e-47;
Matches 55; Conservative 44; Mismatches 94; Indels 12; Gaps 12;

Db 199 ECGELELPQHV-LMNCSDHPLGFSNOCSEFCTDGYQVNPBSEKLECLASGFWIN-KP-P 255
QY 2 DCNELPFRNRTILTGMSDQTYPEGTQAIYKRCRGYSGLGVNIWVC-RKGEWALNLR 60
Db 256 QCLAQO-PKLIKIPERNNICLSAKAFHOSSCSCECFALVGP-EVVOCTASGYWT 313
QY 61 KQKRCPCGHGPT-FGRTT-LTGNGVFYGYKAYITCEGYQLLGEINRCDGDP-WT 117
Db 314 APAPYCKAVQCOHLEAPSEGT-DC-VHPLTFAVYGSSCKFCQCPGYRKGIDMLRCDIS 371
QY 118 NDPIQEVVKKLPVTPAPENKGIYVSAMEPDREYHFGQAVFVNCNSGKIEGDEMHCSDD 177
Db 372 GHWASPLPCEALISCEPLESPVHGS 396
QY 178 GFWSKPKPCVEISCKSPDV-INGS 201

RESULT 10

ENTRY #type complete
TITLE apolipoprotein H homolog precursor - vaccinia virus
ALTERNATE_NAMES #formal_name vaccinia virus
ORGANISM host Homo sapiens (man)
DATE 31-Dec-1989 #sequence_revision 30-Jun-1990 #text_change 22-Jun-1999

ACCESSIONS

ACCESSIONS A31005; B42504
#authors Kotwal, G.J.; Moss, B.
#journal Nature (1988) 335:176-178
#title Vaccinia virus encodes a secretory polypeptide structurally related to complement control proteins.

#cross-references M01D:88318974

#accession A31005
#molecule-type DNA
#residues 1-263 #label KOT
#cross-references GB:X1316; NID:960690; PIDN:CAA31564.1; PID:960691
#experimental_source strain WR

REFERENCE

A42501
#authors Goebel, S.J.; Johnson, G.P.; Perkins, M.E.; Davis, S.W.; Winslow, J.P.; Paolletti, E.
#journal Virology (1990) 179:517-563
#title Appendix to "The complete DNA sequence of vaccinia virus".
#accession B42504
#molecule-type DNA
#residues 1-263 #label GOE
#cross-references GB:M35027; NID:9335317; PIDN:AAA47997.1; PID:9335345

#experimental_source strain Copenhagen

REFERENCE

A42531
#authors Goebel, S.J.; Johnson, G.P.; Perkins, M.E.; Davis, S.W.; Winslow, J.P.; Paolletti, E.
#journal Virology (1990) 179:247-266
#title The complete DNA sequence of vaccinia virus.

#cross-references M01D:91021027
#note annotation: possible protein-coding frames neither amino acid nor nucleotide sequence is given
#superfamily herpesvirus complement control protein;
#complement factor H repeat homology
duplication; extracellular protein

KEYWORDS

FEATURE 1-19
20-263
21-81 #domain signal sequence #status predicted #label SIG
86-143 #product C4b-binding protein homolog #status predicted
148-201 #label MAT
206-261 #domain complement factor H repeat homology #label FH1
#domain complement factor H repeat homology #label FH2
#domain complement factor H repeat homology #label FH3
#domain complement factor H repeat homology #label FH4

SUMMARY

#length 263 #molecular-weight 28629 #checksum 8152
Query Match 19.3%; Score 303; DB 1; Length 263;
Best Local Similarity 34.8%; Pred. No. 1.69e-46;
Matches 65; Conservative 28; Mismatches 76; Indels 18; Gaps 14;

Db 40 ANANNIGDITLYLCLPGRKQKMPYAKCTGTG-TL-ENOCIKRCPSPRIDNQ 96
QY 20 SDQYEGPTQAIYKRCRGYS--LGNVINYCRGKGVNANPLKCKCRKCGHGDPTFT 77
Db 97 LDI-GG-V-DFGSSITSCSGYHLIGESKYSCELTSGTSMWNPAPICSEYKCSPPS 153
QY 78 FTLLGGVVFYGYKAYITCEGYQLLGEIN-YRE-CDTGS-WTNDIPICEVVKCIPVTA 133
Db 154 ISNGR-HNGYE-DE-YTDSVVTYSCNSGYSLTNSGYVLCSCGE-WS-DPFCQVVKP 207
QY 134 PENKTIYSAMEPDREYHFGQAVFVNCNSGKIEGDEMHCSDDGFWSKPKPCVEISCK 193
Db 208 HPTISNG 214
QY 194 SPDYING 200

RESULT 11

ENTRY #type complete
TITLE P-selectin precursor - mouse
ALTERNATE_NAMES CD62; granule membrane protein 140; PADGEM
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 10-Sep-1997

ACCESSIONS

ACCESSIONS A42755; A44899
#authors Weller, A.; Isenmann, S.; Vestweber, D.
#journal J. Biol. Chem. (1992) 267:15176-15183
#title Cloning of the mouse endothelial selectins. Expression of both E- and P-selectin is inducible by tumor necrosis factor alpha.

#cross-references M01D:92340571

#accession A42755
#status nucleic acid sequence not shown
#molecule-type mRNA
#residues 1-768 #label WEL
#cross-references GB:M87861; NID:q200552; PID:q200553
#experimental_source endothelial cells
#note Sequence extracted from NCBI backbone (NCBIP:109467)

REFERENCE

A44899
#authors Sanders, W.E.; Wilson, R.W.; Ballantyne, C.M.; Beaudet, A.L.
#journal Blood (1992) 80:795-800
#title Molecular cloning and analysis of in vivo expression of murine P-selectin.
#cross-references M01D:92345617
#accession A44899
#status nucleic acid sequence not shown

cell adhesion; glycoprotein; phosphohistidine; phosphoprotein; transmembrane protein

1-41

```

SUMMARY
#length 768  #molecular-weight 83098  #checksum 2334

```

Query Match	19.38;	Score 303;	DB 2;	Length 768;
Best local similarity:	32.18;			

220 FSNFSCOTFSCAGGYELDGPGL-QCLASGIWNNPPKCDAVQCSLEAPPHGTM-AC-M 276
| : : | | | : : : : | | | : : : : : : : : : :
86 FEYGVAAVYTCEBGYQLLGEINREDCTDG-WTNDIPICEVVKKCPVTAPENGGIVSSAM 144

Db 277 HPAAFAAYDSSCKFECCOPGYRARGSNLTLLHCTGSGQWSEPLPTCEAIACEPPEIPHGS 334
145 EPDREYHFGQAVRFVCNSYKIEGDEMHCSDDGFWSEKPKCYEISCKSPDV-INS 201

ENTRY TITLE	#type complete
153821 P-selectin - rat	

ORGANISM	#formal_name	Rattus norvegicus	#common_name	Norway rat
DATE	29-May-1998	#sequence_revision	29-May-1998	#text_change
ACCESSTIONS	13-Aug-1999			
FE0301				

#authors Auchampach, J.A.; Oliver, M.G.; Anderson, D.C.; Manning, A.M.
#journal Gene (1994) 145:251-255

#title	Cloning, sequence comparison and in vivo expression of the gene encoding rat P-selectin.
#accession	T53821
#cross-references	MUID:9433817

```
##molecule_type mRNA
##residues 1-768 ##label RES
```

```
#cross-references GB:L23088; NID:9349552; PID:9349553
CLASSIFICATION #superfamily complement factor H repeat homology: C-type
                 lectin homology; EGF homology
FEATURE
```

262-319	#domain EGF homology #label EGF\
E10	#domain complement factor H repeat homo]or
E11	

	#domain	complement	factor	H repeat	homology	#label	FN
510-567	#domain	complement	factor	H repeat	homology	#label	FN105
580-637	#domain	complement	factor	H repeat	homology	#label	FN06
642-659	#domain	complement	factor	H repeat	homology	#label	FN07
669-699	#domain	complement	factor	H repeat	homology	#label	FN08
#length	768	#molecule	1	1	1	1	

Query Match	Score	DB 2;	Length
19.38;	303;	2;	768;

Local similarity 29.3%; Pred. No. 1.69e-46;
Matches 49; Conservative 33; Mismatches 75; Indels 10. Game 10

RESULT	13	
ENTRY		
NBMS		
#type complete		

TITLE	ALTERNATE_NAMES
apolipoprotein H precursor - mouse	50K serum glycoprotein; activated

ORGANISM	beta-2-glycoprotein I	activated protein C-binding protein
DATE	#formal_name Mus musculus #common_name house mouse	
	31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change	

REFERENCE	A43286
#authors	Nonaka, M.; Matsuda, Y.; Shiroishi, T.; Moriyoshi, T.

M.; Natsunume-Sakai, S.
Genomics (1992) 13:1082-1087
Molecular cloning of mouse beta-2-glycoprotein I and mapping
of the gene to chromosome 11.
#cross-references MIMD:92372000
#accession A43286

```
##cross-references GB:D10056
##note the authors translated the
```

#note	REFERENCE
Sequence extracted from NCBI backbone (NCBI:111791, NCBI:111794) and corrected to correspond with the nucleotide translation	JC2243

#journal	Biochem. Biophys. Res. Commun. (1994)	200:1521-1528
#title	Characterization of	

beta2-glycoprotein I (apolipoprotein H).
#cross-references MWID:94242017
#accession JC2243

```
#cross-references GB:S70439; MID:g546780; PIDN:AAB30789.1; PID:g546781
#experimental source liver
```

```
GENETICS
#gene B2gp1
#map_position 11
#accession #11
```

KEYWORDS
chylomicron; duplication; glycoprotein; HDL; heparin binding
lipid binding; monomer; plasma; water

FEATURE	1-19	20-345	35-70
#domain	signal	sequence	#status predicted #label str\
#product	apolipoprotein H	#status predicted #label MAT\	

142-200	#domain	complement	factor	H repeat	homology	#label	FH3\
205-260	#domain	complement	factor	H repeat	homology	#label	FH3\

	#domain complement factor H repeat	homology	#label FH4	#label FH5
264-325				
23-66,51-79,84-124,				
110-137,142-188,				
174-200,205-248,				
234-260,264-315,				
300-307,325-345				
105,117,162,183,				
193				
#disulfide_bonds	#status predicted\			
#binding_site carbohydrate (asn)	(covariant)			#status

SUMMARY #length 345 #molecular-weight 38618 #checksum 2016

Query Match 18.9%; Score 297; DB 1; Length 345;
Best Local Similarity 29.0%; Pred. No. 4,55e-45;
Matches 51; Conservative 37; Mismatches 78; Indels 10; Gaps 8;

Db 40 SYDPGQIYYSCKPGVSGRMRFTCLPLTGMW-PINTLR-CVPRVCPAGLLENGIVRY 97
QY 23 TYPEGQAIYKCRPGVRSIGNVIM-VCR-KGEMVALNPLKRCQKPCGHGDPGTFCTL 80

Db 98 TS---FEYFNKISFACNPGF-FLNGTSSCKTEEGKMSPDIPACARITCPRPPVFKALL 153
QY 81 TGGNVFEYGVKAVYTCNDEGYOLLGEINRECDTDG-WTNDIPICEVYKCLPYTAPBENGKI 139

Db 154 KQYRPSAGNNSLYQDTVYKCLPHFAMINDVYMCTEGQNWRL-PECEVVKCPFP 208
QY 140 VSSAMEPDEHYEGQAVRFVCSNGYKIEGDEMHCSDDGFWSKRKPCEVISCSP 195

RESULT 14

ENTRY A35046 #type complete
TITLE endothelial leukocyte adhesion molecule 1 precursor - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 27-Jul-1990 #sequence_revision 27-Jul-1990 #text_change 16-Jul-1999

ACCESSIONS A38615; A35046
REFERENCE A38615
#authors COLLINS, T.; WILLIAMS, A.; JOHNSTON, G.I.; KIM, J.; EDDY, R.; SHOWS, T.; GIMBRONE JR., M.A.; BEVILACQUA, M.P.
#journal J. Biol. Chem. (1991) 266:2466-2473
#title Structure and chromosomal location of the gene for endothelial-leukocyte adhesion molecule 1.
#cross-references MIM:91115870
#accession A38615
#status preliminary
#molecule_type DNA
#residues 1-610 #label COL
#cross-references GB:M61893; GB:M58017; NID:g182043; PID:g182046
REFERENCE A35046
#authors HESSION, C.; OSBORN, L.; GOFF, D.; CHI-ROSSO, G.; VASSALLO, C.; PASEK, M.; PITTLACK, C.; TIARD, R.; GOELZ, S.; MCCARTHY, K.; HOPPLE, S.; LOBB, R.
#journal Proc. Natl. Acad. Sci. U.S.A. (1990) 87:1673-1677
#title Endothelial leukocyte adhesion molecule 1: direct expression cloning and functional interactions.
#cross-references MIM:90175359
#accession A35046
#status preliminary
#molecule_type mRNA
#residues 1-610 #label HES
#cross-references GB:M30640; NID:g182047; PID:g182048
CLASSIFICATION #superfamily C-type lectin homology; complement factor H repeat homology
KEYWORDS transmembrane protein
FEATURE
12-138 #domain C-type lectin homology #label LCN\
180-237 #domain complement factor H repeat homology #label FH01\
242-299 #domain complement factor H repeat homology #label FH02\
304-362 #domain complement factor H repeat homology #label FH03\
367-425 #domain complement factor H repeat homology #label FH04\
430-488 #domain complement factor H repeat homology #label FH05\
493-547 #domain complement factor H repeat homology #label FH06
SUMMARY #length 610 #molecular-weight 66655 #checksum 542

Query Match 18.9%; Score 298; DB 2; Length 610;
Best Local Similarity 27.4%; Pred. No. 2,63e-45;
Matches 45; Conservative 33; Mismatches 79; Indels 7; Gaps 6;

Db 210 CDGGLPSSMETMOCSSGWSA--PIPCNVVECDVATNPANGFVECFONPGSEFPWNTT 267
QY 34 CRPGYRSLGNVIMCRK-GEWVALNPLKRCQKRPCHGDPGTFGT-LTGNVFEYGVK 91

Db 268 CTFDECEGEFLMAOSL-OCTSSGNWMDNEKPICKATVCRVAPQONGS-VRCSSPAGEF 325
QY 92 AVYTCTNGEYOLLGEINRECDTDG-WTNDIPICEVYKCLPYTAPBENGKIVSSAMEPDEHY 150

Db 326 TFKSSCNFTCEEGFMLOGPAVECTTQOGWTOQIPVCEAFQCTA 369
QY 151 HFGQAVRFVCSNGYKIEGDEMHCSDDGFWSKRKPCEVISCSP 194

Search completed: Thu Jun 8 21:37:54 2000
Job time : 17 secs.

Db 268 CTFDECEGEFLMAOSL-OCTSSGNWMDNEKPICKATVCRVAPQONGS-VRCSSPAGEF 325
QY 92 AVYTCTNGEYOLLGEINRECDTDG-WTNDIPICEVYKCLPYTAPBENGKIVSSAMEPDEHY 150

Db 326 TFKSSCNFTCEEGFMLOGPAVECTTQOGWTOQIPVCEAFQCTA 369
QY 151 HFGQAVRFVCSNGYKIEGDEMHCSDDGFWSKRKPCEVISCSP 194

RESULT 15

ENTRY A32606 #type complete
TITLE endothelial leukocyte adhesion molecule precursor - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999

ACCESSIONS A32606
REFERENCE A32606
#authors BEVILACQUA, M.P.; STENGELIN, S.; GIMBRONE JR., M.A.; SEED, B. Science (1989) 243:1160-1165
#journal Endothelial leukocyte adhesion molecule 1: an inducible receptor for neutrophils related to complement regulatory proteins and lectins.
#title
#cross-references MIM:89162047
#accession A32606
#molecule_type mRNA
#residues 1-610 #label BEV
#cross-references GB:M24736; NID:g537523; PID:g537524
CLASSIFICATION #superfamily EGF homology; complement factor H repeat homology
KEYWORDS transmembrane protein
FEATURE
1-21 #domain signal sequence #status predicted #label SIG\
22-610 #product endothelial leukocyte adhesion molecule 1
#status predicted #label MAT\
143-174 #domain EGF homology #label EGF\
180-237 #domain complement factor H repeat homology #label FH07\
242-299 #domain complement factor H repeat homology #label FH08\
304-362 #domain complement factor H repeat homology #label FH09\
367-425 #domain complement factor H repeat homology #label FH10\
430-488 #domain complement factor H repeat homology #label FH11\
493-547 #domain complement factor H repeat homology #label FH12
SUMMARY #length 610 #molecular-weight 66681 #checksum 746

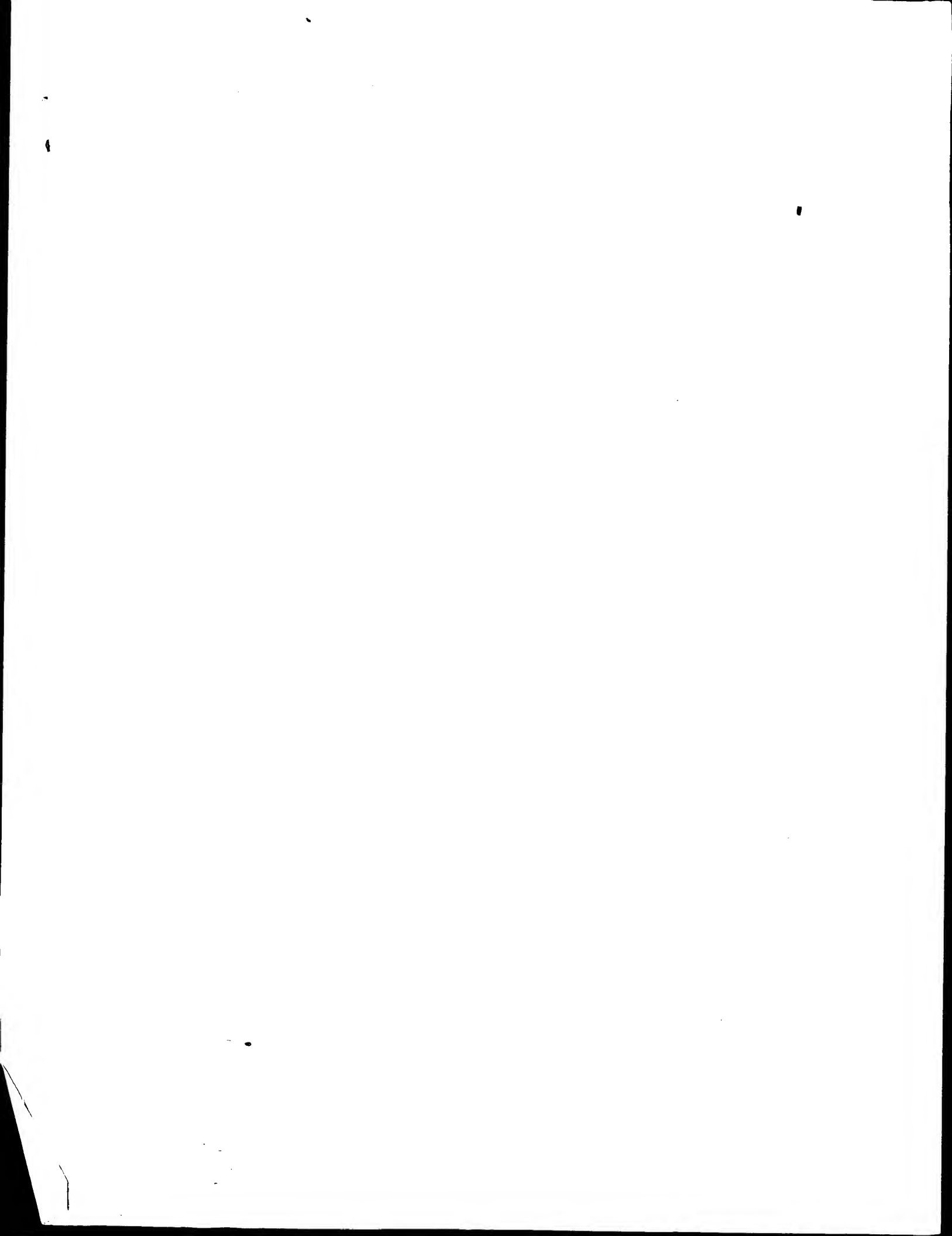
Query Match 18.9%; Score 298; DB 2; Length 610;
Best Local Similarity 27.4%; Pred. No. 2,63e-45;
Matches 45; Conservative 33; Mismatches 79; Indels 7; Gaps 6;

Db 210 CDGGLPSSMETMOCSSGWSA--PIPCNVVECDVATNPANGFVECFONPGSEFPWNTT 267
QY 34 CRPGYRSLGNVIMCRK-GEWVALNPLKRCQKRPCHGDPGTFGT-LTGNVFEYGVK 91

Db 268 CTFDECEGEFLMAOSL-OCTSSGNWMDNEKPICKATVCRVAPQONGS-VRCSSPAGEF 325
QY 92 AVYTCTNGEYOLLGEINRECDTDG-WTNDIPICEVYKCLPYTAPBENGKIVSSAMEPDEHY 150

Db 326 TFKSSCNFTCEEGFMLOGPAVECTTQOGWTOQIPVCEAFQCTA 369
QY 151 HFGQAVRFVCSNGYKIEGDEMHCSDDGFWSKRKPCEVISCSP 194

Search completed: Thu Jun 8 21:37:54 2000
Job time : 17 secs.



 WIRE RELEASE

 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Merch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jun 8 21:38:42 2000; MasPar time 21.93 Seconds
 654.600 Million cell updates/sec

Tabular output not generated.

Title: >US-09-316-163-9

Description: (1-207) from US09316163.pep

Perfect Score: 1573

Sequence: 1 EDCNELPPRRNTEILTSMS.....VEISCKSPDYINGSPISQKI 207

Scoring table: PAM 150
 Gap 11

Searched: 225878 seqs, 69334122 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database:

sptarchb12
 1:sp.archaea 2:sp.bacteria 3:sp.fungi 4:sp.human
 5:sp.invertebrate 6:sp.mammal 7:sp.mnc 8:sp.orcanelle
 9:sp.phage 10:sp.plant 11:sp rodent 12:sp.unclassified
 13:sp.vertebrate 14:sp.virus

Statistics: Mean 41.889; Variance 59.758; scale 0.701

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1573	100.0	449	014570	COMPLEMENT FACTOR H PR	0.00e+00
2	679	43.2	669	028085	CCP MODULES 3-12, WITH	4.01e-156
3	503	32.0	1053	091275	COMPLEMENT REGULATORY	3.14e-106
4	364	23.1	645	09WRU2	COMPLEMENT BINDING PRO	6.97e-68
5	344	21.9	550	040912	ORF 04.	1.73e-62
6	343	21.8	550	040912	ORF 04.	3.22e-62
7	320	20.3	360	091708	COMPLEMENT CONTROL PRO	4.56e-56
8	316	20.1	259	14	41KBP FRAGMENT FROM LE	5.28e-55
9	312	19.8	754	028290	CELL ADHESION MOLECULE	6.08e-55
10	310	19.7	560	023228	COSMID T07H6.	2.06e-53
11	308	19.6	482	028982	E-SELECTIN.	6.98e-53
12	304	19.3	657	014006	COMPLEMENT H FACTOR (F	7.96e-52
13	297	18.9	740	095508	DJ780M13.1.2 (SELECTIN	5.58e-50
14	295	18.8	347	4	DECAV-ACCELERATION FAC	1.87e-49
15	293	18.6	533	11	ACROSOMAL MATRIX COMPO	6.29e-49
16	290	18.4	417	11	512 ANTIGEN (FRAGMENT)	3.85e-48
17	290	18.4	559	11	COMPLEMENT REGULATORY	3.85e-48
18	288	18.3	722	4	DJ780M13.1.1 (SELECTIN	1.29e-47
19	286	18.2	240	4	MEMBRANE COFACTOR PROT	4.30e-47
20	284	18.1	363	6	PORCINE MEMBRANE COFAC	1.43e-46

RESULT ID	1	PRELIMINARY:	PRT:	449 AA.
AC	014570; P78435;			
DT	01-NOV-1996 (TRENBLER, 01, Created)			
DT	01-NOV-1996 (TRENBLER, 01, Last sequence update)			
DT	01-NOV-1999 (TRENBLER, 12, Last annotation update)			
DE	COMPLEMENT FACTOR H PRECURSOR.			
GN	HF OR CFH.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
OC	Eutheria; Primates; Catarrhini; Homidae; Homo.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE: 86134059.			
RT	RIPOLCHE J., DAY A.J., HARRIS T.J.R., SIM R.B.: "The complete amino-acid sequence of human complement factor H."; J. Biochem., 249:593-602(1988).			
RL	[2]			
RN	SEQUENCE OF 226-449 FROM N.A.			
RP	MEDLINE: 86169701.			
RA	KRISTENSEN T., WETSEL R.A., TACK B.F.: "Structural analysis of human complement protein H: homology with C4b binding protein, beta 2-glycoprotein I, and the Ba fragment of B2."; J. Immunol. 136:3407-3411(1986).			
RL	[3]			
RP	SEQUENCE OF 1-19 FROM N.A.			
RA	VIT D.P., WILLIAMS S.A.: Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.			
RL	[4]			
RP	SEQUENCE OF 1-9 FROM N.A.			
RA	DOMINGUEZ O.: Thesis (1993), Immunologia, Hospital Trias I Pujol, Spain.			
RL	EMBL: X07523; CA30403.1.			
DR	EMBL: M12383; AAA52013.1.			
DR	EMBL: U56979; AAB01987.1.			
DR	EMBL: Z29665; CAA82763.1.			
DR	HSSP: P10998; IYVC.			
DR	PRAM: PF00084; sushi; 7.			
FT	Signal.			
KW	Signal.			
FT	CHAIN			
FT	SEQUENCE			

ALIGNMENTS

Query Match 100.0%; Score 1573; DB 4; Length 449;

Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 19 EDCNLEPPRRNTELLTGSMSOTVPEGTQAIYKCRPGRSIGNIYMCRCRKEWALNPLR 78
QY 1 EDCNLEPPRRNTELLTGSMSOTVPEGTQAIYKCRPGRSIGNIYMCRCRKEWALNPLR 60
Db 79 KCQRRPGHGDPTPGFTLLTGNGVFEYGVAVYCNNGYOLLGEINRECDTDGWTNDI 138
QY 61 KCQRRPGHGDPTPGFTLLTGNGVFEYGVAVYCNNGYOLLGEINRECDTDGWTNDI 120
Db 139 PICEVVKLLPTAPENGKIVSAMPDEHYHFGQAVRVCNSGYKIEGDEEMHCSDDGF 198
QY 121 PICEVVKLLPTAPENGKIVSAMPDEHYHFGQAVRVCNSGYKIEGDEEMHCSDDGF 180
Db 199 SKERKCVIEISCKSPDVINGSPISOKI 225
QY 181 SKERKCVIEISCKSPDVINGSPISOKI 207

RESULT 2
ID Q28085 PRELIMINARY; PRT: 669 AA.

AC Q28085;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
DE CCP MODULES 3-12, WITH PARTS OF CCP 2 AND 13 (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;
RN Bos; Bos.
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE; 96202005.
RA SOAMES C.J., DAY A.J., SIM R.B.;
RT "Prediction from sequence comparisons of residues of factor H involved in the interaction with complement component C3b."
RL Biochem. J. 315:523-531(1996).
DR EMBL; X98697; CAA67257.1; -.
DR HSSP; P10998; 1VVC.
PRFAM; PF00084; sush1. 11.
FT NON_TER 1
FT NON_TER 669
SQ SEQUENCE 669 AA; 75683 MW; FAFOD174 CRC32;

Query Match 43.2%; Score 679; DB 6; Length 669;
Best Local Similarity 66.1%; Pred. No. 4.01e-156;
Matches 84; Conservative 21; Mismatches 22; Indels 0; Gaps 0;

Db 7 LAGNDFEYCAKYVYTDDEGYOVGMNPRECDTNGWTNDIPICEVVKCLPVTPEPENGKI 66
QY 80 LAGNDFEYCAKYVYTDDEGYOVGMNPRECDTNGWTNDIPICEVVKCLPVTPEPENGKI 139
Db 67 FSPALDEDDYTYGGVAVFECNSGYMDGPKQIHCAGSAGSALTEPKCEIFCKPVIYN 126
QY 140 VSSAMPEDEHYHFGQAVRVCNSGYKIEGDEEMHCSDDGFWSKERKCVIEISCKSPDVIN 199
Db 127 GOAVLRK 133
QY 200 GSPISOK 206

RESULT 3
ID Q91275 PRELIMINARY; PRT: 1053 AA.

AC Q91275;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
DE COMPLEMENT REGULATORY PLASMA PROTEIN.
OS Paratubax nebulifer (barred sand bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;

OC Perciformes; Percoidae; Serranidae; Paratubax.

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE; 94318039.
RA DAHMEN A., KAI DOH T., ZIPPEL P.F., GIGLI I.;
RT "Cloning and characterization of a cDNA representing a putative complement-regulatory plasma protein from barred sand bass (Paratubax nebulifer)."
RT Biochem. J. 301:391-397(1994).
DR EMBL; U21703; AAA92556.1; -.
DR HSSP; P08603; 1HPH.
PRFAM; PF00084; sush1. 16.
SQ SEQUENCE 1053 AA; 117597 MW; 0D68EDB CRC32;

Query Match 32.0%; Score 503; DB 13; Length 1053;
Best Local Similarity 41.8%; Pred. No. 3.14e-106;
Matches 76; Conservative 29; Mismatches 68; Indels 9; Gaps 7;

Db 47 EASYGGQVRVGVGV-S-GFEKLVCEGKMETRGA--KCQRRSGHGEDAOFADPHL 102
QY 21 DQTFEGTQAIYKCRPGRSIGNIYMCRCRKEWALNPLRKCQRRSGHGEDTPTGFTLL 80
Db 103 AEGNDFVGSKVYTCQKGYOMVSRINRCVAGCMGVVPCVSCQC-PATHDNVVOY 161
QY 81 TCGNVEYGVAVYCNNGYOLLGEINRECDTDGWTNDIPICEVVKCLPVTAPENGKIV 140
Db 162 TGG--PE-EATFGVAVRPSCKSRSEIILGSPELYCDERGMVSGVPCKAITCAIPPIEN 218
QY 141 SSAMPEDEHYHFGQAVRVCNSGYKIEGDEEMHCSDDGFWSKERKCVIEISCKSPDVIN 199
Db 219 GN 220
QY 200 GS 201

RESULT 4
ID Q9WRU2 PRELIMINARY; PRT: 645 AA.

AC Q9WRU2;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
DE COMPLEMENT BINDING PROTEIN.
OS Macaca mulatta rhadinovirus 17577.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE; 99174001.
RA SEARLES R.P., BERGOUAN E.P., AXTHELM M.K., WONG S.W.;
RT "Sequence and genomic analysis of a Rhesus macaque rhadinovirus with similarity to Kaposi's sarcoma-associated herpesvirus/human herpesvirus 8."
RT J. Virol. 73:3040-3053(1999).
DR EMBL; AF083501; AAD21332.1; -.
SQ SEQUENCE 645 AA; 71526 MW; 60FB82D6 CRC32;

Query Match 23.1%; Score 364; DB 14; Length 645;
Best Local Similarity 32.6%; Pred. No. 6.97e-68;
Matches 61; Conservative 40; Mismatches 71; Indels 15; Gaps 12;

Db 337 EKYVGSASVELICRPGTKMSTVSECLSNGTWTPANA--KCHRRKCPDPELLNGEYI 394
QY 22 QTYPEGTQAIYKCRPGRSIGNIY-C-RKGEWALNPLRKCQRRSGHGEDTPTGFT 79
Db 395 VTSGEDAFKYGTNITKCNNGYOLLGSNWRICMKRDLTVMEKAPICDIECKKPPQ 454
QY 80 LTGG-INFEYGVAVYTCNDEGTOLLGE-IN-Y--RCC-DTDMWTNDIPICEVVKCLPVT 133
Db 455 ITNGKY--HEVK-DF-YOYLDVTTFSCNDRFSLVSGDEMTTCISNT-WNKPFRCOITS 509
QY 134 PENGKIVSAMPEDRYHFGQAVRVCNSGYKIEGDEEMHCSDDGFWSKERKCVIEISCK 193

Db	510	APNIAHG	516	
Qy	194	SPDYING	200	
RESULT	5	PRELIMINARY;	PRT;	550 AA.
ID	040912			
AC	040912;			
DT	01-JAN-1998 (TREMBlrel. 05, Created)			
DT	01-JAN-1998 (TREMBlrel. 05, Last sequence update)			
DT	01-NOV-1999 (TREMBlrel. 12, Last annotation update)			
DE	ORF 04.			
OS	Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Gammaherpesvirinae; Rhadinovirus.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 97296220.			
RA	NEIPEL F., ALBRECHT J.C., FLECKENSTEIN B.;			
RT	"Cell-homologous genes in the Kaposi's sarcoma-associated rhadinovirus			
RT	human herpesvirus 8: determinants of its pathogenicity?";			
RL	J. Virol. 71:4187-4192(1997).			
DR	EMBL; U93872; AAB62602.1; -.			
DR	HSSP; P10998; 1YVC.			
DR	PFAM; PF00084; sush1; 4.			
SQ	SEQUENCE 550 AA; 60648 MW; 558098B4 CRC32;			
Query Match	21.9%;	Score 344;	DB 14;	Length 550;
Best Local Similarity	33.9%;	Pred. No. 1,73e-62;		
Matches	61;	Conservative 38;	Mismatches 64;	Indels 17; Gaps 13
Db	52	RCRSGITTYARNITATCAGGCTW-S-EPTATCNKSKCPNPEIONGKIYFHGGDALKYG	109	
Qy	33	KCRPGYRSLG- NVIMWC- RKGEVALNLPKRCQKRCCHPDPTFGPTTLTG- NVEYG	89	
Db	110	ANISVNCNGYFLVGRVYRCMIGASQGMWSSSPCEKEKC-H-R-P--KIENGDF	163	
Qy	90	VKAYTTCNEGIIOLG-E-INTRECTDG---WTNDIPICEVYKCLPYAPENGRKIVSSAM	144	
Db	164	RKDKDYEVNDVAHEECNEGTYLLVPHSIIACAVANNVTWSMPTCEIACCKFPSTHTGYPI	223	
Qy	145	EPDREYH-FGCAVREVCNSGYKIEEDDEMHCSDDGFSMEKPKVCVEISCKSPDYINGSPI	203	
RESULT	6	PRELIMINARY;	PRT;	550 AA.
ID	P88903			
AC	P88903;			
DT	01-MAY-1997 (TREMBlrel. 03, Created)			
DT	01-MAY-1997 (TREMBlrel. 03, Last sequence update)			
DT	01-NOV-1999 (TREMBlrel. 12, Last annotation update)			
DE	ORF 4.			
OS	Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Gammaherpesvirinae; Rhadinovirus.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 97121480.			
RA	RUSO J.J., BOHEWZY R.A., CHIEN M.C., CHEN J., YAN M., MADDALENA D.,			
RA	PARRY J.P., PERUZZI D., EDELMAN I.S., CHANG Y., MOORE P.S.;			
RT	"Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus			
RT	(HHV8)";			
RL	Proc. Natl. Acad. Sci. U.S.A. 93:14862-14867(1996).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 97094384.			
RA	MOORE P.S., BASHOFF C., WEISS R.A., CHANG Y.;			
RT	"Molecular miniv. of human cytokine and cytokine response pathway			
RT	genes by KSHV";			
RL	Science 274:1739-1744(1996).			
DR	EMBL; U75698; AAC57082.1; -.			
DR	HSSP; P10998; 1YVC.			
DR	PFAM; PF00084; sush1; 4.			
SQ	SEQUENCE 550 AA; 60687 MW; CE2A2CAEC CRC32;			

Query Match	21.8%	Score 343	DB 14	Length 550
Best Local Similarity	33.9%	Pred. No. 3.22e-62		
Matches	61	Conservative	38	Mismatches 64; Indels 17; Gaps 13;
Db	52	RCRSGYTYANINATLOGGTW-S-EPTATCNKKSQCPNPGETIONKGVIPHGODALKYG	109	
Qy	33	KCRGRGYSLSG-NVIMWC-RKGEWALNPLRKQCRKQRPFGHDPFGFTLTGG-NWFEYRG	89	
Db	110	ANISYVNEGFLVGRGVYRCMIGASGQAMSSSPFCCKENC-H-R-P---KIKKXGDF	163	
Qy	90	VKAYATTCNEGVLGG-E-INTRECDTGG---WINDIPICEVWCKLPYTAENGKIVSSAM	144	
Db	164	KPDYQYENYNAVHEFCNEGVTGIVGPHSIACAANNVTSMNPTCELAGCFPSYTHSGPI	223	
Qy	145	EPDRFHY-FGQAVAFVCSGKIKIGDEMHCSDDGFFMSKEXKPKVELISCKSPDINSPI	203	
RESULT	7	PRELIMINARY;	PRT;	360 AA.
ID	09YT08			
AC	09YT08			
DT	01-MAY-1999	(Tremblrel. 10, Created)		
DT	01-MAY-1999	(Tremblrel. 10, Last sequence update)		
DT	01-NOV-1999	(Tremblrel. 12, Last annotation update)		
DE	COMPLEMENT CONTROL PROTEIN HOMOLOG CCPH.			
OS	Atelinae herpesvirus 3.			
OC	Viruses; dsDNA viruses, no RNA stage: Herpesviridae;			
CC	Gammaherpesvirinae; Rhadinovirus.			
RN	SEQUENCE FROM N.A.			
RP	STRAIN-73.			
RA	ALBRECHT J.-C., FLECKENSTEIN B.;			
RT	"Primary structure of the Herpesvirus Atles Genome.";			
RU	Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF083424; AAC95530.1; -.			
DR	HSSP; P10998; IYVC.			
SO	SEQUENCE 360 AA; 40208 MW; F66C7ADA CRC32;			
Query Match	20.3%	Score 320; DB 14; Length 360;		
Best Local Similarity	37.6%	Pred. No. 4.56e-56;		
Matches	71	Conservative 28; Mismatches 72; Indels 18; Gaps 13		
Db	40	SSGSYPNGTLOVTCRKGYIGROIQTVTCVGNM-TV-P-NECQRRCSTPADLINGMT	96	
Qy	20	SDQYTPGSTQAIYKCRPGYRSIGNYIMVCRKGEMWALNPLRKCKRPGHGDPTFGTFT	79	
Db	97	VNG-VLY-VGSVITTCNTMGYOLGSPF-SSCLGPDGRVMMTPRPICETICKKPPETI	153	
Qy	80	LTGGVVEFGKAVATTCNEGFIQLJSEINPRE--DTDG--WINDIPICEVWCKLPYAP	134	
Db	154	ANGT--HTNIK-E-YTYTLDATVATYSCNDETLLTLPSSKOCSEFGRVWPDEETKCEKV	209	
Qy	135	ENGKIVSSAMEPDRRIYHGOAARFCVCSGKYLE--GDEMHCSDDGFFV-SKEXPCVEIS	191	
Db	210	CKIPOVANG 218		
Qy	192	CKSPDIVNG 200		
RESULT	8	PRELIMINARY;	PRT;	259 AA.
ID	P87616			
AC	P87616;			
DT	01-MAY-1997	(Tremblrel. 03, Created)		
DT	01-MAY-1997	(Tremblrel. 03, Last sequence update)		
DT	01-NOV-1999	(Tremblrel. 12, Last annotation update)		
DE	41KBP FRAGMENT FROM LEFT END OF GENOME.			
OS	Coimpo virus (CPV).			
OC	Viruses; dsDNA viruses, no RNA stage: Poxviridae; Chordopoxvirinae;			
CC	Orthopoxvirus.			
RN	SEQUENCE FROM N.A.			
RP	STRAIN-CRI-90;			

RX MEDLINE; 97068532.
 RA SAFRONOV P.F., PETROV N.A., RIAZANKINA O.I., TOTMENIN A.V.,
 RA SHCHELKHNOV S.N., SANDAKHCHIV L.S.,
 RT "Genes of a circle of hosts for the cowpox virus."
 RL Dokl. Akad. Nauk 349:829-833(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GRI-90;
 RX MEDLINE: 98229462.
 RA SHCHELKHNOV S.N., SAFRONOV P.F., TOTMENIN A.V., PETROV N.A.,
 RA RIAZANKINA O.I., GUTOKOV V.V., KOTVAL G.J.;
 RT "The genomic sequence analysis of the left and right species-specific
 terminal region of a cowpox virus strain reveals unique sequences and
 RT a cluster of intact ORFs for immunomodulatory and host range
 RT proteins."
 RL Virology 243:432-460(1998).
 DR EMBL: X94355; CAA64102.1;
 DR EMBL: Y11842; CAA72567.1;
 DR HSSP: P10998; IYVC.
 DR PFAM: PF00084; sush1; 4.
 SQ SEQUENCE 259 AA; 28193 MW; 76531F63 CRC32;

 Query Match 20.1%; Score 316; DB 14; Length 259;
 Best Local Similarity 34.8%; Pred. No. 5,286-55;
 Matches 71; Conservative 30; Mismatches 84; Indels 19; Gaps 15;

 Db 20 CCEIPSRITMKRKGTVDSHYNIGDTIEYLCPGYRKOKMGPITAKCTGTGM-TL-FN 75
 QY 3 CNELPFRNTEILGSMWSDOTYPEGTQAIKCRGYRS--LGNVIMCRGEMVYALPRL 60
 Db 76 QCIKRCPSRDIIDNGOLDI-GG-V-DFGSSITSCNSGYHLIGESKSYCELGSTGSMW 132
 QY 61 KCKRCGHHGDPFCTFTLTLGNVEYGVKAVYTCNEGYYQLLEIN-YRE-CDTGG-W 116
 Db 133 NPAPRCESVCKCSPSISNGR--HNGYE-DF-YTDSVYTYSCNSGYSLIGSGVLCSC 188
 QY 117 TNDIPCEVVKCLPYAPENGKIVSSAMEPDRBYHFGQAVFCVNSGYKLEGDEMHCS 176
 Db 189 GE-WS-DPPTQIQIKCPHPTISNG 210
 QY 177 DGFWSKRPCKVEIICKSPDIVNG 200

 RESULT 9
 ID 028290 PRELIMINARY; PRT; 754 AA.
 AC 028290;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE CELL ADHESION MOLECULE PRECURSOR (FRAGMENT).
 GN GMP140.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-SPLEEN;
 RA MANNING A.M., SANDERS W.E. JR., KUKIELKA G.L., DORE M.,
 RA ROSENBLUM C.L., HAWKINS H.L., MICHAEL L.H., EMTMAN M.L., SMITH C.W.,
 RA BEAUBERT A.L., ANDERSON D.C.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: M88170; AAA63789.1;
 DR HSSP: P16109; IFSB.
 DR PROSITE: PS00615; G_TYPE-LECTIN_1; 1.
 DR PFAM: PF00059; lectin_c; 1.
 DR PFAM: PF00084; sush1; 8.
 KW signal; Cell adhesion; glycoprotein.
 FT NON_TER 1
 FT SIGNAL 1
 FT CHAIN <1 13 POTENTIAL.
 FT CHAIN 14 754 CELL ADHESION MOLECULE.
 SQ SEQUENCE 754 AA; 82303 MW; E9570281 CRC32;

 Query Match 19.8%; Score 312; DB 6; Length 754;

Best Local Similarity 30.5%; Pred. No. 6,086-54;
 Matches 51; Conservative 36; Mismatches 69; Indels 11; Gaps 10;

 Db 212 SFY-CTEGYELNPSKLECLASGTWN-KPPR-CVATQC-PLKTEQSGMNLHSEVAF 267
 QY 30 AIYKCRGYRSLGNVIMVC-RKSEWALNPLRKQCRPGCHPDTP-FGTFT-LTGQNVF 86
 Db 268 QYOSCHFCSEEGFALVGP-EVYQCTASGMMTAAPVCEAVACGPLKSPYGSMDSP-- 324
 QY 87 EYGVKAVYTCNEGYYQLLEINRECDTGG-WTNDIPCEVVKCLPYAPENGKIVSSAME 145
 Db 325 SSAPFYNTSCRHCBGFRLEBADLVQCTDLCQWTAAPACALQC 371
 QY 146 PDREYHFGQAVRFVCSNGYKIEGDEMHCSDDGFWSKRPCKVEIICKSPDIVNG 192

 RESULT 10
 ID 022328 PRELIMINARY; PRT; 560 AA.
 AC 022328;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE COSMID T07H6.
 GN T07H6.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabdilia; Rhabdilitida;
 OC Rhabdilita; Rhabdilitidae; Rhabdilitae; Peloderinae; Caenorhabdilitis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE: 94150718.
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
 RA CRATON M., DEAR S., DU Z., DORBIN R., FAVELLO A., FULTON L.,
 RA GARNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PARSONS J., PERCY C., REIKEN L., ROOPRA A., SANDERS D., SHONKKEEN R.,
 RA SAUNDON N., SMITH A., SONNHAMMER E., STADEN R., STOLTON J.,
 RA THIRRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans."
 RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA GEISEL C.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA WATERSTON R.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U53344; AAA96225.1;
 DR HSSP: P10998; IYVC.
 DR PFAM: PF00084; sush1; 7.
 SQ SEQUENCE 560 AA; 61619 MW; 3A10AF63 CRC32;

 Query Match 19.7%; Score 310; DB 5; Length 560;
 Best Local Similarity 33.6%; Pred. No. 2,066-53;
 Matches 51; Conservative 36; Mismatches 51; Indels 14; Gaps 10;

 Db 117 AQMGPD-LR-CKARACDDPDIDENG---LREGDTFEPHAYKSCNPGFTLVGTS-RQ 170
 QY 51 GEWALNPLRKQCRPGCHPDTPFGFTLLGQNVFEGVAVYTCNEGYYQLLEINRE 110
 Db 171 CSNMGWNEPANKATECSRSPSLHGKVVGSLL-T--YQ-S-VVYSGDHGRVLGQ 224
 QY 111 CDPLGS-WTNDIPCEVVKCLPYAPENGKIVSSAMEPDRBYHFGQAVRFVCSNGIKLEG 169
 Db 225 VQICLAEIGWGNBPCEIRCSVLPTLPNG 256

OY 170 EEMHCSDDGFWSEKPKVCIEISCKS-PDYING 200

```
RESULT 11
ID Q28982 PRELIMINARY; PRT; 482 AA.
AC Q28982;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
DE E-SELECTIN.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97075911.
RA WINKLER H., BROSTJAN C., CSIZMADIA V., NATARAJAN G., ANRATHER J.,
RA BACH F.H.;
RT "The intron-exon structure of the porcine E-selectin-encoding gene.";
RL Gene 176:67-72(1996).
DR EMBL; U37521; AAC48680.1; -.
DR HSSP; P16581; IESL.
DR PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
DR PFAM; PF00008; EGF; 1.
DR PFAM; PF00059; lectin.C; 1.
DR PFAM; PF00084; sushi; 4.
DR PRINTS; PR00343; SELECTIN.
KW Lectin; Glycoprotein.
SQ SEQUENCE 482 AA; 52341 MW; 94528E97 CRC32;
```

Query Match 19.6%; Score 308; DB 6; Length 482;
Best Local Similarity 28.9%; Pred. No. 6, 98e-53;
Matches 52; Conservative 38; Mismatches 82; Indels 8; Gaps 6;

```
Db 196 QSLPWNTCAFECKEFGELLGPEHLOCTSSGSDGKP--TCKAVTCDTVGHPONGDVSC 253
OY 22 QTTEGTQAIYKCRPGYSLGNVIMVC-RKGEWALNLRKCGKRPCHPDPTFGFTL 80
Db 254 NHHSTGEAVYSTCHFTCAEFGLGPAQT-ECTAOGMTQOAPVCAKCPAVSQPKNG 312
OY 81 TGNVFEYGVAAV--WMCNEGYYLLGELINRECDTDG-WINDIPICEVYKCLPYTAPENG 137
Db 313 -LVFTHSPTGEFTYKSCAFSGEFLRGSAQLACTSGOGMTQOEVPSQVYVSCSLEV 371
OY 138 KIVSSAMPEDEHYHFGQAVRVCNSGYKIEGDEHMSDDGFWSEKPKVCIEISCKSPDV 197
```

```
RESULT 12
ID Q14006 PRELIMINARY; PRT; 657 AA.
AC Q14006;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
DE COMPLEMENT H FACTOR (FRAGMENT).
GN HF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 88025472.
RA DAY A.J., RIPOCHE J., LYONS A., MCINTOSH B., HARRIS T.J., SIM R.B.;
RT "Sequence analysis of a cDNA clone encoding the C-terminal end of
RT human complement factor H.";
RL Biosci. Rep. 7:201-207(1987).
DR EMBL; M1517; AAA52016.1; -.
DR HSSP; P08603; 1HFH.
DR PFAM; PF00084; sushi; 11.
FT NON_TER 1
SQ SEQUENCE 657 AA; 74247 MW; F4AB5238 CRC32;
```

Query Match 19.3%; Score 304; DB 4; Length 657;
Best Local Similarity 27.1%; Pred. No. 7, 96e-52;

Matches 56; Conservative 51; Mismatches 86; Indels 14; Gaps 14;

```
Db 296 CSQ-PPQIEHGTINSSRSQSEYAHGKLSYNGEGFRISSEHETTCYMGK-W-S-SP-PQ 351
OY 3 CNELPFRNTEILTGSWSD-TYPEGQAIYKCRPGYSLGNVIMVC-RKGEWALNLRKCGK 61
Db 352 CEGLPCKSPPEISHGVVAHMSD-YOGEERYTKCFEGGIDGPAIAK-CLGKWSHP-P 408
OY 62 CQKRPCHGPDPTFGFTLGNVFEYGVAAVYTCNDEYQVLGEINRECDTDGWTNDIP 121
Db 409 SCIKTDLSLPSENA-IPNGEKK-D-VIKAGEQVYTCATYTKMGDASNVTCINSR-WT 464
OY 122 ICEVYKCLPYTAPENGKIVSSAMEPDREYHFGQAVRVCNSGYKIEGDEHMSDDGFW 181
Db 465 -GRPTCDTSCVNPPTVQNAIVSRQ 490
OY 182 KEKPKVCIEISCKSPDYINGSPISQKI 207
```

Query Match 18.9%; Score 297; DB 4; Length 740;
Best Local Similarity 28.8%; Pred. No. 5, 58e-50;
Matches 51; Conservative 39; Mismatches 75; Indels 12; Gaps 11;

```
Db 281 AFAYSSCKFEQOPGYRVRGLMLRCIDSGHMSA--PLETCAISC-EPLSPVHSGMDC 337
OY 23 TYPEGQAIYKCRPGYSLGNVIMVC-RKGEWALNLRKCGKRPCHPDPTFGFTL 80
Db 338 SPSLRAFOYDINCSEFCAGFMLRGADIVR-CDNLGOWTAPAPVCAQALOCODLPVNEAR 396
OY 81 TGG-INVFEYGVAAVYTCNDEYQVLGEINRECDTDG-WINDIPICEVYKCLPYTAPENG 138
Db 397 -VNCS-HFPGARY-QSVCSFTCNGLLVGASVLOCATGNNMNSVPECAIIPCTP 450
OY 139 IVSSAMEPDREYHFGQAV-RVCNSGYKIEGDEHMSDDGFWSEKPKVCIEISCKS 194
```

```
RESULT 14
ID P78361 PRELIMINARY; PRT; 347 AA.
AC P78361;
DT 01-MAY-1997 (TREMblrel. 03, Created)
DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)
DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
DE DECAY-ACCELERATION FACTOR (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-BRAIN HIPPOCAMPUS;
RA KUMAR V.B., HYUNG C., NAKRA R., WALTERS M., SASSER T., BERNARDO A.;
RL Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
```

 WIRE (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MSrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jun 8 21:38:12 2000; MasPar time 8.92 Seconds
 706.532 Million cell updates/sec

Tabular output not generated.

Title: >US-09-316-163-9
 Description: (1-207) from US09316163.pep
 Perfect Score: 1573
 Sequence: 1 EDCNELPPRRNTETILTSWS.....VEISCKSPDVINGSPISOKI 207

Scoring table: PAM 150
 Gap 11

Searched: 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0
 Listing first 45 summaries

Database: swiss-prot38
 1:swissprot

Statistics: Mean 42.607; Variance 59.776; scale 0.713

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	1573	100.0	1231	1	CFAB_HUMAN COMPLEMENT FACTOR H PR	0.00e+00
2	1154	73.4	1234	1	CFAB_MOUSE COMPLEMENT FACTOR H PR	6.87e-299
3	308	19.6	484	1	LEM2_PIG E-SELECTIN PRECURSOR (8.32e-54
4	308	19.6	830	1	LEM3_HUMAN P-SELECTIN PRECURSOR (8.32e-54
5	303	19.3	263	1	VCP_VACCV P-SELECTIN PRECURSOR (1.82e-52
6	303	19.3	768	1	LEM3_MOUSE P-SELECTIN PRECURSOR (1.82e-52
7	303	19.3	768	1	LEM3_RAT P-SELECTIN PRECURSOR (1.82e-52
8	297	18.9	345	1	APDH_MOUSE BETA-2-GLYCOPROTEIN I	7.31e-51
9	297	18.9	610	1	LEM2_HUMAN E-SELECTIN PRECURSOR (3.95e-51
10	292	18.6	345	1	APDH_BOVIN BETA-2-GLYCOPROTEIN I	1.57e-49
11	293	18.6	381	1	DAF_HUMAN COMPLEMENT DECAV-ACCEL	5.35e-49
12	290	18.4	597	1	C4BP_HUMAN C4B-BINDING PROTEIN AL	5.35e-49
13	288	18.3	360	1	CCPH_HSVSA COMPLEMENT CONTROL PRO	1.82e-48
14	288	18.3	612	1	LEM2_MOUSE E-SELECTIN PRECURSOR (1.82e-48
15	286	18.2	549	1	LEM2_MOUSE E-SELECTIN PRECURSOR (6.17e-48
16	287	18.2	551	1	LEM2_RAT E-SELECTIN PRECURSOR (3.35e-48
17	284	18.1	340	1	DAF_PONPY COMPLEMENT DECAV-ACCEL	2.09e-47
18	280	17.9	377	1	MCP_HUMAN MEMBRANE COFACTOR PROT	1.30e-46
19	280	17.8	507	1	DAF_CAVPO COMPLEMENT DECAV-ACCEL	2.39e-46
20	279	17.7	611	1	LEM2_CANFA E-SELECTIN PRECURSOR (4.40e-46
21	276	17.5	1019	1	LFC_TACTR LIMOUS CLOTTING FACTO	2.72e-45
22	272	17.3	958	1	HIG_PROMO LOCOCOMOTION-RELATED PRO	3.07e-44
23	270	17.2	345	1	APDH_HUMAN BETA-2-GLYCOPROTEIN I	1.03e-43

24	270	17.2	610	1	C4BP_BOVIN	C4B-BINDING PROTEIN AL	1.03e-43
25	270	17.2	1033	1	CR2_HUMAN	COMPLEMENT RECEPTOR TY	6.31e-43
26	267	17.0	2039	1	CR1_HUMAN	COMPLEMENT RECEPTOR TY	1.15e-42
27	266	16.9	485	1	LEM2_BOVIN	E-SELECTIN PRECURSOR (3.85e-42
28	264	16.8	345	1	APDH_CANFA	BETA-2-GLYCOPROTEIN I	2.11e-42
29	265	16.8	459	1	C4BP_MOUSE	C4B-BINDING PROTEIN PR	2.47e-41
30	260	16.5	558	1	C4BP_RAT	C4B-BINDING PROTEIN AL	2.58e-40
31	257	16.3	1025	1	CR2_MOUSE	COMPLEMENT RECEPTOR TY	1.56e-39
32	254	16.1	390	1	DAF1_MOUSE	COMPLEMENT DECAV-ACCEL	9.33e-39
33	251	15.0	769	1	LEM3_SHEEP	P-SELECTIN PRECURSOR (2.09e-35
34	238	15.1	407	1	DAF2_MOUSE	COMPLEMENT DECAV-ACCEL	6.80e-35
35	236	15.0	297	1	APDH_RAT	BETA-2-GLYCOPROTEIN I	3.98e-34
36	233	14.8	646	1	LEM3_BOVIN	P-SELECTIN PRECURSOR (7.15e-34
37	232	14.7	372	1	LEM1_RAT	L-SELECTIN PRECURSOR (1.29e-33
38	231	14.7	372	1	LEM1_MOUSE	L-SELECTIN PRECURSOR (7.89e-31
39	220	14.0	372	1	LEM1_PANTR	L-SELECTIN PRECURSOR (7.89e-31
40	220	14.0	372	1	LEM1_MACQU	L-SELECTIN PRECURSOR (7.89e-31
41	220	14.0	372	1	LEM1_PAPPA	L-SELECTIN PRECURSOR (7.89e-31
42	220	14.0	372	1	LEM1_HUMAN	L-SELECTIN PRECURSOR (2.51e-30
43	218	13.9	372	1	LEM1_PONPY	L-SELECTIN PRECURSOR (7.84e-27
44	204	13.0	661	1	F13B_HUMAN	COAGULATION FACTOR XII	4.17e-25
45	197	12.5	330	1	CFH1_HUMAN	COMPLEMENT FACTOR H-LI	4.17e-25

CFAB_HUMAN STANDARD; PRT; 1231 AA.

01-AUG-1988 (Rel. 08, Created)

01-JAN-1990 (Rel. 13, Last sequence update)

15-FEB-2000 (Rel. 39, Last annotation update)

COMPLEMENT FACTOR H PRECURSOR.

HEF OR HE OR CFH.

Human sapiens (Human).

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

[1]

SEQUENCE FROM N.A.

RC TISSUE-LIVER;

RA MEDLINE; 88134059.

RA Ripoché J., Day A.J., Harris T.J.R., Sim R.B.;

RT "the complete amino acid sequence of human complement factor H.";

RL Biochem. J. 249:593-602(1988).

[2]

SEQUENCE OF 53-445 FROM N.A.

RA MEDLINE; 87054207.

RA Schulz T.F., Schwaeble W., Stanley K.K., Weiss E., Dierich M.P.;

RT "Human complement factor H: isolation of cDNA clones and partial cDNA

sequence of the 38-kDa tryptic fragment containing the binding site

for C3b.";

RL Eur. J. Immunol. 16:1351-1355(1986).

[3]

SEQUENCE OF 226-445 FROM N.A. AND PARTIAL SEQUENCE.

RA MEDLINE; 86169701.

RA Kristensen T., Wetzel R.A., Tack B.F.;

RT "Structural analysis of human complement protein H: homology with C4b

binding protein, beta-2-glycoprotein I, and the Ba fragment of B2.";

RL J. Immunol. 136:3407-3411(1986).

[4]

SEQUENCE OF 1047-1231 FROM N.A.

RA MEDLINE; 91201892.

RA Estaller C., Koistinen V., Schwaeble W., Dierich M.P., Weiss E.H.;

RT "Cloning of the 1.4-kb mRNA species of human complement factor H

reveals a novel member of the short consensus repeat family related

to the carboxy terminal of the classical 150-kDa molecule.";

RL J. Immunol. 146:3190-3196(1991).

[5]

SEQUENCE OF 19-35.

RA MEDLINE; 83048213.

RA Sim R.B., Discipio R.G.;

RT "Purification and structural studies on the complement-system control

protein beta 1H (factor H).";

[illegible]

SUCH AS HEPARIN, PHOSPHOLIPIDS, AND DEXTRAN SULFATE, MAY PREVENT ACTIVATION OF THE INTRINSIC BLOOD COAGULATION CASCADE BY BINDING TO PHOSPHOLIPIDS ON THE SURFACE OF DAMAGED CELLS.

-1- TISSUE SPECIFICITY: PLASMA.

-1- SIMILARITY: CONTAINS 4 SUSHI (SCR) REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb.sib.ch).

CC EMBL: D10056; BAA00945.1; -

DR EMBL: S70439; AAB30789.1; -

DR EMBL: Y11356; CAA72190.1; -

DR PIR: A43286; NBSM.

DR HSSP: P10998; 1VVC.

DR MGD: MGI:88058; APOH.

DR PFAM: PF00084; sushi; 4.

KW Heparin-binding; glycoprotein; Plasma; Repeat; Sushi; Signal.

FT SIGNAL 1 19

FT CHAIN 20 345

FT DOMAIN 22 261

FT REPEAT 22 80

FT REPEAT 83 138

FT REPEAT 141 201

FT REPEAT 204 261

FT DOMAIN 263 345

FT DISULFID 23 66

FT DISULFID 51 79

FT DISULFID 84 124

FT DISULFID 110 137

FT DISULFID 142 188

FT DISULFID 174 200

FT DISULFID 205 240

FT DISULFID 234 260

FT DISULFID 264 315

FT DISULFID 300 325

FT DISULFID 307 345

FT CARBOHYD 162 162

FT CARBOHYD 183 183

FT CARBOHYD 193 193

FT CONFLICT 252 252

SEQ SEQUENCE 345 AA; 38619 MW; C83F8A6BBD51C940 CRC64;

Query Match 18.9%; Score 297; DB 1; Length 345;

Best Local Similarity 29.0%; Pred. No. 7,31e-51;

Matches 51; Conservative 37; Mismatches 78; Indels 10; Gaps 8;

DB 40 SYDPEQIVYCKRGVGRGMRRFTCLTGMW--PINTLR-CVPRVCPFAGILENGIVRY 97

QY 23 TYPESTQIVYCRPEYRGLGVIM-VCR-KGEWALNLRKQCRKPGCHPDDTFPGFTL 80

DB 98 TS--FEFKNISFACNGF-FLNCTSSSKCTEGKWSPDIPACARITCPPPVKFAALL 153

QY 81 TGGNFEFGVAVYTCNGYQLGEINVRCDTDLG-WTNDIPICGVYVCLPVTAPENGI 139

DB 154 KDIPSAGNSIYQDTVYFKCLPFAMAGNDVYACTEGGNTRL-PELEVKCFPP 208

QY 140 VSSAMEPREYHFGAVRFVNSGYKIEGDEMHCSDDGFWSKERPKCVETLSCSP 195

RESULT 9

ID LEM2_HUMAN STANDARD: PRT: 610 AA.

AC P16581; P16111; Created)

DI 01-APR-1990 (Rel. 14, Last sequence update)

DI 01-AUG-1990 (Rel. 15, Last annotation update)

DI 15-JUL-1999 (Rel. 38, Last annotation update)

DE E-SELECTIN PRECURSOR (ENDOTHELIAL LEUKOCYTE ADHESION MOLECULE 1)

DE (ELAM-1) (LEUKOCYTE-ENDOTHELIAL CELL ADHESION MOLECULE 2) (LECAM2)

DE (CD62E).

GN SELE OR ELAM1.

OS Homo sapiens (Human).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

CC Eutheria; Primates; Catarrhini; Hominoidea; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE: 90175359.

RA Hession C., Osborn L., Goff D., Chi-Rosso G., Vassallo C.,

RA Pasek M., Pittack C., Tizard R., Goetz S., McCarthy K., Hopple S.,

RA Lobb R.;

RT "Endothelial leukocyte adhesion molecule 1: direct expression cloning and functional interactions."

RL Proc. Natl. Acad. Sci. U.S.A. 87:1673-1677(1990).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE: 89162047.

RA Bevilacqua M.P., Stengelin S., Gimbrone M.A. Jr., Seed B.;

RT "Endothelial leukocyte adhesion molecule 1: an inducible receptor for neutrophils related to complement regulatory proteins and lectins."

RL Science 243:1160-1165(1989).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE: 91115870.

RA Collins T., Williams A., Johnston G.I., Kim J., Eddy R., Shows T.,

RA Gimbrone M.A. Jr., Bevilacqua M.P.;

RT "Structure and chromosomal location of the gene for endothelial-leukocyte adhesion molecule 1."

RL J. Biol. Chem. 266:2466-2473(1991).

RN [4]

RP LIGAND.

RX MEDLINE: 91068005.

RA Phillips M.L., Nudelma E., Gaeta F.C., Perez M., Singhal A.K.,

RA Hakomori S., Paulson J.C.;

RT "ELAM-1 mediates cell adhesion by recognition of a carbohydrate ligand, sialyl-Lex."

RL Science 250:1130-1132(1990).

RN [5]

RP 3D-STRUCTURE MODELING OF LECTIN DOMAIN.

RX MEDLINE: 93202275.

RA Mills A.;

RT "Modelling the carbohydrate recognition domain of human E-selectin."

RL FEBS Lett. 319:5-11(1993).

RN [6]

RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 22-178.

RX MEDLINE: 94150646.

RA Graves B.J., Crowther R.L., Chandran C., Rumberger J.M., Li S.,

RA Huang K.-S., Presky D.H., Familletti P.C., Wolitzky B.A., Burns D.K.;

RT "Insight into E-selectin/ligand interaction from the crystal structure and mutagenesis of the lec/EGF domains."

RL Nature 367:532-538(1994).

RN [7]

RP VARIANT ARG-149.

RX MEDLINE: 95179107.

RA Wenzel K., Felix S., Kleber F.X., Brachold R., Menke T., Schatke S.,

RA Schulte K.L., Glaser C., Rohde K., Baumann G., Speer A.;

RT "E-selectin polymorphism and atherosclerosis: an association study."

RL Hum. Mol. Genet. 3:1935-1937(1994).

RN [8]

RP VARIANT ARG-149.

RX MEDLINE: 99134508.

RA Ye S.Q., Usher D., Virgil D., Zhang L.Q., Yochim S.E., Gupta R.;

RT "A beta polymorphism detects the mutation of serine-128 to arginine in CD 62 gene - a risk factor for coronary artery disease."

RL J. Biomed. Sci. 6:18-21(1999).

CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF POLYACROSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF GLYCOPOLYIDS).

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC -1- POLYMORPHISM: A POLYMORPHISM IN POSITION 149 IS ASSOCIATED WITH A HIGHER RISK OF CORONARY ARTERY DISEASE (CAD). A SIGNIFICANTLY HIGHER MUTATION FREQUENCY (ARG-149) IS OBSERVED IN PATIENTS WITH

"Molecular basis of reduced or absent expression of RT decay-accelerating factor in Cromer blood group phenotypes.";

RT Blood 84:1276-1282(1994).

CC -1- FUNCTION: THIS PROTEIN RECOGNIZES C4B AND C3B FRAGMENTS THAT CC CONDENSE WITH CELL-SURFACE HYDROXYL OR AMINO GROUPS WHEN NASCENT CC C4B AND C3B ARE LOCALLY GENERATED DURING C4 AND C3 ACTIVATION. CC INTERACTION OF DAF WITH CELL-ASSOCIATED C4B AND C3 POLYPEPTIDES CC INTERFERES WITH THEIR ABILITY TO CATALYZE THE CONVERSION OF C2 AND CC FACTOR B TO ENZYMATICALLY ACTIVE C2A AND Bb AND THEREBY PREVENTS OF CC THE FORMATION OF C4B2A AND C3BBB, THE AMPLIFICATION CONVERSAES OF CC THE COMPLEMENT CASCADE.

CC -1- FUNCTION: ALSO ACTS AS THE RECEPTOR FOR ECHOVIRUS 7 AND RELATED CC VIRUSES (ECHOVIRUSES 13, 21, 29 AND 33).

CC -1- SUBUNIT: MONOMER (MAJOR FORM) AND NON-DISULFIDE-LINKED, COVALENT CC HOMODIMER (MINOR FORM).

CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR. CC DAF-1) ARE PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE. CC -1- TISSUE SPECIFICITY: EXPRESSED ON THE SURFACES OF EPITHELIAL CELLS CC TYPES THAT ARE IN INTIMATE CONTACT WITH PLASMA MEMBRANES OF ALL CELL CC PROTEINS. IT IS ALSO FOUND ON THE SURFACES OF EPITHELIAL CELLS CC Lining EXTRACELLULAR COMPARTMENTS, AND VARIANTS OF THE MOLECULE CC ARE PRESENT IN BODY FLUIDS AND IN EXTRACELLULAR MATRIX.

CC -1- DOMAIN: THE FIRST SUSHI DOMAIN (SCR1) IS NOT NECESSARY FOR CC FUNCTION. SCR2 AND SCR4 PROVIDE THE PROPER CONFORMATION FOR THE CC ACTIVE SITE ON SCR3 (BY SIMILARITY).

CC -1- PTM: THE SER/THR-RICH DOMAIN IS HEAVILY O-GLYCOSYLATED. CC -1- POLYMORPHISM: DAF IS RESPONSIBLE FOR THE CROMER BLOOD GROUP CC SYSTEM. IT CONSISTS OF AT LEAST SEVEN HIGH-INCIDENCE (CR(A), CC TC(A), DR(A), ES(A), WES(B), UMC, AND IFC) AND LOW-INCIDENCE CC (TC(B), TC(C), AND WES(A)) ANTIGENS THAT RESIDE ON DAF. IN THE CC CROMER PHENOTYPES DR(A-) AND INAB THERE IS REDUCED OR ABSENT CC EXPRESSION OF DAF, RESPECTIVELY. IN THE CASE OF THE DR(A-) CC PHENOTYPE, A SINGLE NUCLEOTIDE SUBSTITUTION WITHIN EXON 5 ACCOUNTS CC FOR TWO CHANGES: A SIMPLE AMINO ACID SUBSTITUTION THAT IS THE CC BASIS OF THE ANTIGENIC VARIATION, AND AN ALTERNATIVE SPLICING CC EVENT THAT UNDERLIES THE DECREASED EXPRESSION OF DAF IN THIS CC PHENOTYPE.

CC -1- SIMILARITY: CONTAINS 4 SUSHI (SCR) REPEATS. CC -1- SIMILARITY: BELONGS TO THE RECEPTORS OF COMPLEMENT ACTIVATION CC (RCA) FAMILY.

CC -1- DATABASE: NAME-PROW; NOTE-CD guide CD55 entry; CC WWW="http://www.ncbi.nlm.nih.gov/prow/CD/cd55.htm"

CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration CC between the Swiss Institute of Bioinformatics and the EMBL outstation CC the European Bioinformatics Institute. There are no restrictions on its CC use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See http://www.isb-sib.ch/announce/ CC or send an email to license@isb-sib.ch).

DR EMBL: M31516; AAA52169.1; -

DR EMBL: M30142; AAA52168.1; -

DR EMBL: M15799; AAA52167.1; -

DR EMBL: M64653; AAA52170.1; -

DR EMBL: M64356; AAA52170.1; JOINED.

DR EMBL: S72858; AAC60633.1; -

DR PIR: B26359; B26359.

DR PIR: A26359; A26359.

DR PIR: S16187; S16187.

DR PIR: A39101; A39101.

DR PIR: S23138; S23138.

DR HSSP: P08603; IHCC.

DR MIM: 125240; -

DR PFAM: PF00084; sush1.4.

KM Complement pathway; Plasma; glycoprotein; Membrane; Repeat;

KW Alternative splicing; GPI-anchor; Signal; Polymorphism;

FT SIGNAL 1 34

FT CHAIN 35 353 COMPLEMENT DECAY-ACCELERATING FACTOR.

FT PROPEP 354 381 REMOVED IN MATURE FORM.

FT DOMAIN 284 4 X SUSHI (SCR) REPEATS.

FT REPEAT 35 95 SUSHI 1.

FT REPEAT 97 159 SUSHI 2.

FT REPEAT 162 221 SUSHI 3.

FT REPEAT 224 284 SUSHI 4.

FT DOMAIN 287 356 SER/THR-RICH.

FT DISULFID 36 81

FT DISULFID 65 94

FT DISULFID 98 145

FT DISULFID 129 158

FT DISULFID 163 204

FT DISULFID 190 220

FT DISULFID 225 267

FT DISULFID 253 283

FT CARBOHYD 95 95

FT LIPID 353 353

FT VARIANT 52 52

FT VARIANT 52 52

FT VARIANT 82 82

FT VARIANT 199 199

FT VARIANT 227 227

FT VARSPLIC 362 381

FT CONFLICT 80 80

FT CONFLICT 85 85

FT SEQUENCE 381 AA: 41388 MW: 29138EBE64B565E CRC64;

Query Match 18.6%; Score 293; DB 1; Length 381;

Best Local Similarity 29.7%; Pred. No. 8; 51e-50;

Matches 38; Conservative 43; Mismatches 75; Indels 19; Gaps 14;

Db 98 C-EVPTRLNSASIKOPYIIQNTFPVGVVEYECRPYRREPSLSPRLTGLQMLKSTAYE 156

QY 3 CNELPRRMTTELITGWSMOTY-PECTOAIYKCRGRIYSLGAVI-MVC-KRGEWVALNP 58

Db 157 F-CRKSQSPNNGEIRNGQIDVPGC-IL-FGATISFSCNTGYKLFSGTSSFCLLISGSVQ 212

QY 59 LKRCQRPGHGDPPFGFTLTGNGVEYGVKAVYTCNEGQLGELIN-YR-ECDTIG- 115

Db 213 WSDPLEECREIYCPAPQIDNG-IIQG-ERDH-YGYNQSVYIAANKGPTMGESITCT 268

QY 116 WINDPILICVVCIVTAPENKIVSNAMEPDRYHFGQAVRFVCGNSGKIEGDEMHC 175

Db 269 VNNDGEMSGPPPEC 283

QY 176 ---DDGFWSEKPKC 187

RESULT 12

ID C4BP_HUMAN STANDARD; PRT; 597 AA.

AC P04003;

DT 23-OCT-1986 (Rel. 02, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE C4B-BINDING PROTEIN ALPHA CHAIN PRECURSOR (PROLINE-RICH PROTEIN)

GN C4BPA OR C4BP.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

CC Eutheria; Primates; Catarrhini; Homiidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-LIVER;

RX MEDLINE; 90073699.

RA Matsunuchi T., Okamura S., Aso T., Sata T., Niho Y.;

RT "Molecular cloning of the cDNA coding for proline-rich protein (PRP);

RT Biochem. Biophys. Res. Commun. 163:138-144(1989).

OY 135 ENKIVSSAMEPDREYHFGAVRFVNSGYKIEGDEMHCS-D-D--GFWSEKPKVEI 190
 Db 238 TCRKPDVSHGEMVS 251
 : : : : :
 OY 191 SCKSPDVINGSPIS 204

RESULT 13
 ID CCPH.HSVSA STANDARD; PRT; 360 AA.
 AC 001016;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE COMPLEMENT CONTROL PROTEIN HOMOLOG PRECURSOR (CCPH).
 GN 4 OR CCPH.
 OS Herpesvirus saimiri (strain 11).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 92333688.
 RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,
 RA Newman C., Wiltman S., Craxton M.A., Coleman H., Fleckenstein B.,
 RA Honess R.W.;
 RT "Primary structure of the herpesvirus saimiri genome."
 RL J. Virol. 66:5047-5058(1992).
 RN [2]
 RP SIMILARITY TO CCP.
 RX MEDLINE: 92260674.
 RA Albrecht J.-C., Fleckenstein B.;
 RT "New member of the multigene family of complement control proteins in
 RT herpesvirus saimiri."
 RL J. Virol. 66:3937-3940(1992).
 CC -1- ALTERNATIVE PRODUCTS: A MEMBRANE-BOUND FORM AND A SECRETED FORM
 CC ARE PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE.
 CC -1- SIMILARITY: BELONGS TO THE SUPERFAMILY OF THE REGULATORS OF
 CC COMPLEMENT ACTIVATION (RCA).
 CC -1- SIMILARITY: CONTAINS 3 SUSHI (SCR) REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X64346; CAA45626.1; -
 DR EMBL: X64346; CAA45627.1; -
 DR EMBL: X60283; CAA42823.1; -
 DR EMBL: X60283; CAA42822.1; -
 DR PIR: BA2534; WMBE2E.
 DR PIR: BA2534; WMBE2E.
 DR PIR: S24567; S24567.
 DR HSSP: P10998; IIVC.
 DR PIR: PF00084; SUSHI; 4.
 DR PIR: PF00084; SUSHI; 4.
 DR Signal; Repeat; SUSHI; Transmembrane; Alternative splicing;
 DR Glycoprotein.
 DR SIGNAL
 FT CHAIN 1 20 POTENTIAL.
 FT DOMAIN 83 265 COMPLEMENT CONTROL PROTEIN HOMOLOG.
 FT REPEAT 83 143 SUSHI 1.
 FT REPEAT 146 206 SUSHI 2.
 FT REPEAT 209 265 SUSHI 3.
 FT TRANSMEM 328 350 POTENTIAL.
 FT DISULFID 84 125 BY SIMILARITY.
 FT DISULFID 111 142 BY SIMILARITY.
 FT DISULFID 147 191 BY SIMILARITY.
 FT DISULFID 175 205 BY SIMILARITY.
 FT DISULFID 210 252 BY SIMILARITY.
 FT DISULFID 238 264 BY SIMILARITY.
 FT CARBOHYD 36 36 POTENTIAL.
 FT CARBOHYD 39 39 POTENTIAL.

FT CARBOHYD 46 46 POTENTIAL.
 FT CARBOHYD 72 72 POTENTIAL.
 FT CARBOHYD 155 155 POTENTIAL.
 FT CARBOHYD 294 294 POTENTIAL.
 FT VARSPIC 289 302 RINGNCTSMPTQ -> AECACGSGNYPIS (IN
 FT VARSPIC 289 302 SHORT ISOFORM).
 FT VARSPIC 303 360 MISSING (IN SHORT ISOFORM).
 SQ SEQUENCE 360 AA; 40006 MW; 6278A6C2ECD49669 CRC64;

Query Match 18.3% Score 288; DB 1; Length 360;
 Best Local Similarity 34.68; Pred. No. 1,82e-48;
 Matches 64; Conservative 33; Mismatches 70; Indels 18; Gaps 13;

Db 44 YPNGTLLHVTGREGYARPPQVITCVGNM-TV-P-KKCKKCSFPPDNLNGRYVTG- 99
 OY 24 YPSTQALTYKCRPGRYSLGNVIMVCKRGKGVNLPKCKRCGCGHGDTPFSTFTLTGG 83
 Db 100 NLY-YGCVITYTNGSYSLIGSTT-SACLKRGKRVMTPRPICDKCKP--PQIAN 155
 OY 84 NVEYGVKAVYTCNEGQYLGEINREC--DTDG--TWNDIPICEVVKCLPTADENGR 138
 Db 156 GTHDNV-K-DE-YTTLDTVYMSCNDEKLTITGSSKLCSTGSGWVNGETKCEFIKCLP 213
 OY 139 IVSSAMEPDREYHFGAVRFVNSGYKIE--GDEMHCSDDGFW-SKEKPKVEISKSP 195
 Db 214 QVANA 218
 OY 196 DVIWG 200

RESULT 14
 ID LEM2.MOUSE STANDARD; PRT; 612 AA.
 AC 000690;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE E-SELECTIN PRECURSOR (ENDOTHELIAL LEUKOCYTE ADHESION MOLECULE 1)
 DE (ELAM-1) (LEUKOCYTE-ENDOTHELIAL CELL ADHESION MOLECULE 2) (LECAM2)
 DE (CD62E).
 GN SELE OR ELAM-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 92283265.
 RA Becker-Andre M., van Huijsduijn R.H., Losberger C., Whelan J.,
 RA Delamarier J.F.;
 RT "Murine endothelial leukocyte-adhesion molecule 1 is a close
 RT structural and functional homologue of the human protein.";
 RL Eur. J. Biochem. 206:401-411(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 92340571.
 RA Weller A., Isenmann S., Vestweber D.;
 RT "Cloning of the mouse endothelial selectins. Expression of both E-
 RT and P-selectin is inducible by tumor necrosis factor alpha.";
 RL J. Biol. Chem. 267:15176-15183(1992).
 CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
 CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
 CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
 CC GLYCOPOLIDS).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN
 CC -1- SIMILARITY: CONTAINS 6 SUSHI (SCR) REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M80778; AAA37547.1; -
 CC EMBL; M87862; AAA37577.1; ALT_INIT.
 DR HSSP; P16581; IKJA.
 DR MGD; MGI:98378; SELE.
 DR PRINTS; PRO0343; SELECTIN.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
 DR PROSITE; PS50041; C-TYPE_LECTIN_2; 1.
 DR PFAM; PF00008; EGF; 1.
 DR PFAM; PF00059; lectin_C; 1.
 DR PFAM; PF00084; sushi; 6.
 DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
 KM Selectin; Signal; Sushi; Repeat.
 FT CHAIN 1 21
 FT SIGNAL 1 21
 FT CHAIN 22 612
 FT TRANSMEM 22 557
 FT DOMAIN 558 579
 FT DOMAIN 580 612
 FT DOMAIN 139 175
 FT DOMAIN 179 549
 FT REPEAT 179 239
 FT REPEAT 242 301
 FT REPEAT 304 364
 FT REPEAT 367 427
 FT REPEAT 430 490
 FT REPEAT 493 549
 FT DISULFID 40 138
 FT DISULFID 111 130
 FT DISULFID 143 134
 FT DISULFID 148 163
 FT DISULFID 165 174
 FT DISULFID 180 225
 FT DISULFID 210 238
 FT DISULFID 243 287
 FT DISULFID 273 300
 FT DISULFID 305 350
 FT DISULFID 336 363
 FT DISULFID 368 413
 FT DISULFID 399 426
 FT DISULFID 431 476
 FT DISULFID 462 489
 FT DISULFID 494 535
 FT DISULFID 521 548
 FT CARBOHYD 25 25
 FT CARBOHYD 145 145
 FT CARBOHYD 192 192
 FT CARBOHYD 203 203
 FT CARBOHYD 266 266
 FT CARBOHYD 333 333
 FT CARBOHYD 391 391
 FT CARBOHYD 504 504
 FT CARBOHYD 528 528
 FT SEQUENCE 612 AA; 66749 MW; 86F05713F0EC23D CRC64;

Query Match 18.38; Score 288; DB 1; Length 612;
 Best Local Similarity 25.18; Pired. No. 1,82e-48;
 Matches 49; Conservative 42; Mismatches 95; Indels 9; Gaps 8;

Db 242 ECEBAL-HPAHGRKSSNGSYPAWTTCTPDCEGVRVGAONLQCTSSGIMDNETP-- 298
 Oy 2 DCELPRLPRRTTELLTGSWSDOTPECTOAIYKRPYRISGNIYWC-RKGEVVALNPLR 60
 Db 299 SCRAVTCDAIPDPONGFVSCSHSTAGELAKSSCNFTCEOSFTLGGPAOV-ECSAOGONT 357
 Oy 61 KCQKRPQGH-P-GDTPGFTLTGNGVFEYGVAVYTCNCGYOLLEINRECDTDG-WT 117
 Db 358 POLPYCAVCEALSAQGNMFKLPSASGP-FONGSSCFSESEFELKGSRLCCGPR 416

Qy 118 NDIPCEVVKCLPYAPAPNGKIVSSAMEPDRHYFGQAVRVCSGIKIBGDEMHCSDD 177
 Db 417 GEMDSKRPCTSAVKC 431
 Qy 178 GWSKREKRCVHISC 192

RESULT 15
 ID LEM2_RAT STANDARD; PRT; 549 AA.
 AC P98105;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE E-SELECTIN PRECURSOR (ENDOTHELIAL LEUKOCYTE ADHESION MOLECULE 1)
 DE (ELAM-1) (LEUKOCYTE-ENDOTHELIAL CELL ADHESION MOLECULE 2) (LECAM2)
 DE (CD62E).
 GN SELE OR ELAM-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG;
 RA Rosenthal C.L., Auchampach J.A., Anderson D.C., Manning A.M.;
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases
 CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
 CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
 CC ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF
 CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
 CC GLYCOPOLIPIDS)
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMS.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 5 SUSHI (SCR) REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; L28527; AAA41113.1; -
 DR HSSP; P16581; IKJA.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
 DR PROSITE; PS50041; C-TYPE_LECTIN_2; 1.
 DR PFAM; PF00008; EGF; 1.
 DR PFAM; PF00059; lectin_C; 1.
 DR PFAM; PF00084; sushi; 5.
 DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
 KW Selectin; Signal; Sushi; Repeat.
 FT CHAIN 1 21
 FT SIGNAL 1 21
 FT CHAIN 22 549
 FT TRANSMEM 22 494
 FT DOMAIN 495 516
 FT DOMAIN 517 549
 FT DOMAIN 139 138
 FT DOMAIN 179 175
 FT DOMAIN 179 486
 FT REPEAT 179 239
 FT REPEAT 242 301
 FT REPEAT 304 364
 FT REPEAT 367 427
 FT REPEAT 430 486
 FT DISULFID 40 138
 FT DISULFID 111 130
 FT DISULFID 143 134
 FT DISULFID 148 163

```

FT DISULFID 165 174 BY SIMILARITY.
FT DISULFID 180 225 BY SIMILARITY.
FT DISULFID 210 238 BY SIMILARITY.
FT DISULFID 243 287 BY SIMILARITY.
FT DISULFID 273 300 BY SIMILARITY.
FT DISULFID 305 350 BY SIMILARITY.
FT DISULFID 336 363 BY SIMILARITY.
FT DISULFID 368 413 BY SIMILARITY.
FT DISULFID 399 426 BY SIMILARITY.
FT DISULFID 431 472 BY SIMILARITY.
FT DISULFID 458 485 BY SIMILARITY.
FT CARBOHYD 25 25 POTENTIAL.
FT CARBOHYD 60 60 POTENTIAL.
FT CARBOHYD 145 146 POTENTIAL.
FT CARBOHYD 192 192 POTENTIAL.
FT CARBOHYD 203 203 POTENTIAL.
FT CARBOHYD 266 266 POTENTIAL.
FT CARBOHYD 313 313 POTENTIAL.
FT CARBOHYD 320 330 POTENTIAL.
FT CARBOHYD 333 333 POTENTIAL.
FT CARBOHYD 441 441 POTENTIAL.
FT CARBOHYD 465 465 POTENTIAL.
SQ SEQUENCE 549 AA: 60079 MW: 85CECD7B0144C8 CRC64;

```

Query Match 18.2% Score 286; DB 1; Length 549;
 Best Local Similarity 25.1%; Pred. No. 6,17e-48;
 Matches 50; Conservative 45; Mismatches 97; Indels 7; Gaps 7;

```

DB 242 ECKALTQPAHG-VKCCSSNPGSYPMWTTCTFDCEGRRYGAONLOCTSSGWDNEKPSG 300
OY 2 DNELEPPRRNTEILGSMDSMDQTYPEGTQAIKCRPGYRSIGNYIMWC-RKGEWVALNPLR 60
DB 301 KAVTCDAIRPONGSVSCSNSTAGAL-AFKSSCNFTCEHSTFLQSPAQV-ECSAQGQWTP 358
OY 61 KQOK-RPCGHPGDTFRGFTTLTGANVFEGYKAVYTCNEGYYDLGEINXRECDTDG-WTN 118
DB 359 QIPVCKASOCALSAPOQGHMKCLPSASAP-FQSGSSCKFSCDEGFELKGSRRLOCGPRG 417
OY 119 DIPICEVVKCLFVTAPENGKIYVSAMEPDRREYHFGQAVRFVCSGKIIGDEEMHCSDDG 178
DB 418 EWDSEKPTCAGVQCSLDL 436
OY 179 FWSKEKPKCVETSCSKSPDV 197

```

Search completed: Thu Jun 8 21:38:24 2000
 Job time : 12 secs.


```

Mpsrch_pp    protein - protein database search, using Smith-Waterman algorithm
Run on:      Thu Jun  8 21:37:08 2000;      MasPar time 8.38 Seconds
Tabular output not generated.               585.300 Million cell updates/sec

```

Scoring table: PAM 150
Gap 11

Post-processing: Minimum Match 08
Listing first 45 summaries

Database: a-geneseq36

Statistics: Mean 31.247; Variance 114.255; scale 0.273

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	973	61.9	240	1	M39154	Human partial Complement	6.80e-97
2	842	53.5	216	1	M39155	Clone pBR95FHA10 CFH r	8.32e-82
3	308	19.6	464	1	W99892	Porcine E-selectin.	1.19e-21
4	308	19.6	830	1	R65216	P-selectin.	1.19e-21
5	298	18.9	610	1	R08116	Endothelial cell-leuco	1.43e-20
6	298	18.9	610	1	R05494	Endothelial leukocyte	1.43e-20
7	298	18.9	610	1	R05538	Endothelial-leukocyte	1.43e-20
8	298	18.9	610	1	W18839	E-selectin.	1.43e-20
9	298	18.9	610	1	W46731	Amino acid sequence of	1.43e-20
10	293	18.6	299	1	W06881	Decay accelerating fac	4.96e-20
11	293	18.6	381	1	W26317	Human decay accelerat	4.96e-20
12	293	18.6	381	1	R66683	Decay accelerating fac	4.96e-20
13	293	18.6	381	1	P70048	Human decay accelerat	4.96e-20
14	293	18.6	381	1	P94773	Decay accelerating fac	4.96e-20
15	293	18.6	381	1	W73505	Decay accelerating fac	4.96e-20
16	293	18.6	440	1	P94783	Membrane bound decay	4.96e-20
17	293	18.6	440	1	P27483	Human glycoposphatidy	4.96e-20
18	293	18.6	440	1	P70049	Human decay accelerat	4.96e-20
19	293	18.6	440	1	R66684	Membrane accelerating fac	4.96e-20
20	293	18.6	577	1	W06882	Membrane co-factor pro	1.04e-19
21	290	18.4	581	1	R13490	Human C4 binding prote	1.71e-19
22	288	18.3	302	1	W26320	Hepesvirus secreted g	1.71e-19
23	288	18.3	302	1	R55793	Hepesvirus salmatri sc	1.71e-19

24	288	18.3	360	1	W65319	1.71e-19
25	288	18.3	360	1	HE5792	1.71e-19
26	286	18.2	263	1	P92203	2.81e-19
27	284	18.1	363	1	W12414	4.60e-19
28	281	17.9	376	1	R93939	9.66e-19
29	281	17.9	377	1	W27484	9.66e-19
30	281	17.9	377	1	R93941	9.66e-19
31	281	17.9	384	1	CD46 construct <i>s</i> hBSCR3	9.66e-19
32	281	17.9	384	1	R86319	9.66e-19
33	278	17.7	324	1	R10924	2.03e-18
34	278	17.7	373	1	R15233	2.03e-18
35	278	17.7	373	1	R15229	2.03e-18
36	278	17.7	377	1	CD46 from clone pm5.6	3.32e-18
37	276	17.5	646	1	Porcine P-selectin.	1.58e-18
38	276	17.5	251	1	W14945	3.32e-18
39	276	17.5	254	1	R13361	3.32e-18
40	276	17.5	279	1	Membrane cofactor prot	3.32e-18
41	276	17.5	279	1	W65360	3.32e-18
42	276	17.5	357	1	rsMCP protein SMO ID N	4.25e-18
43	275	17.5	377	1	R10927	4.25e-18
44	275	17.5	378	1	Human membrane cofactor	4.25e-18
45	273	17.4	350	1	CD46 construct subSCR2	6.95e-18
					CD46 construct subSCR2	6.95e-18
					CD46 from clone pm5.3	6.95e-18
					CD46 from clone pm5.10	6.95e-18

ALIGNMENTS

RESULT	1
ID	W39154 standard; Protein; 240 AA.

27-APR-1998 (first entry)
DE Human partial Complement factor H protein fragment 1.
KW Complement factor H: tumour associated antigen; renal cancer;
KW urogenital cancer; medicament; modulator.
OS Homo sapiens.
PN WC9738136-Al.
PD 16-OCR-1997.
PT 09-APR-1997; U05710.
PR 06-MAR-1997; US-812481.
PR 09-APR-1996; US-015083.
PR 09-APR-1996; US-610048.
PR 06-MAR-1997; US-038614.
PA (BARO-) BARO DIAGNOSTIC SCI INC.
PI Enfield DL, Hass GM, Kinders RJ;
WP1: 97-512742/47.
DR N-PSDB: V02790.
PT Treating or screening for cancer, e.g. renal or urogenital cancer -
PT by modulating or detecting tumour associated human complement Factor
PT H related antigen, or nucleic acid encoding it
PS Example 6B: Fig 6B, 104pp; English.
CC This partial protein sequence represents a region of the human
CC tumour associated complement factor H (CFH). This sequence is used
CC in the identification of complement factor H related proteins and
CC antigens isolated from clone PBB95FH10 (see W39155). The detection of
CC such proteins and a CFH antigens can be used in screening or for the
CC treatment of renal or urogenital cancer, e.g. bladder, cervical or
CC prostate cancer. Agents that may modulate this antigen could be used in
CC the manufacture of a medicament for the treatment of a tumour cell.
SQ Sequence 240 AA;

Query Match	61.9%;	Score 973;	DB 1;	Length 240;
Best Local Similarity	100.0%;	Pred. No. 6,80e-97;		
Matches 130; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

```
Db      1 FTLTGNVEFGYKAYVTCNTEGTYOLLGEINRCECTDGMNDIPICEVVKCLPTAPANG   60
        |||||
Qy      78 FTLTGNVNEFGYKAYVTCNTEGTYOLLGEINRCECTDGMNDIPICEVVKCLPTAPENG   137
        |||||

Db      61 KIYSAMEPDRREHYFQGAVFVCNSGYKIETGDEMHCSDDGFWSKERKPVEICKSPDV   120
        |||||
Qy      138 KIYSAMEPDRREHYFQGAVFVCNSGYKIETGDEMHCSDDGFWSKERKPVEICKSPDV   197
        |||||

Db      121 INGSPIISQIKI 130
        |||||
```

QY 198 INGSPIISOKI 207

RESULT 2

ID W39155 standard; Protein; 216 AA.

AC W39155; (first entry)

DT 27-APR-1998

DE Clone PRB9FH410 CFH related protein fragment.

KW Complement factor H; tumour associated antigen; renal cancer; urogenital cancer; modulator.

OS Synthetic.

PN W09738136-A1.

PD 16-OCT-1997.

PF 09-APR-1997; 005710.

PR 06-MAR-1997; US-812481.

PR 09-APR-1996; US-015083.

PR 06-MAR-1997; US-038614.

PA (BARD) BARD DIAGNOSTIC SCI INC.

PI Enfield DL, Haas GM, Kinders RJ;

DR WPI; 97-512742/47.

DR N-PSDB; V02791.

PT Treating or screening for cancer, e.g. renal or urogenital cancer -

PT H related antigen, or nucleic acid encoding it

PS Example 6B; Fig 6B; 104pp; English.

CC This partial protein is found in clone PRB9FH410 and represents a

CC complement factor H related protein with homology to a region of the

CC human tumour-associated complement factor H (CFH). The detection of this

CC protein and a CFH antigen can be used in screening for or the treatment

CC of renal or urogenital cancer, e.g. bladder, cervical or prostate cancer.

CC Agents that may modulate this antigen could be used in the manufacture of

CC a medicament for the treatment of a tumour cell.

SQ Sequence 216 AA.

Query Match Best Local Similarity 53.5%; Score 842; DB 1; Length 216;

Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 TCNEGYQLGELINRECDTGTNDIPICVYKCLPTAPANGKTVSSAMEPDRYHFGQ 60

QY 95 TCNEGYQLGELINRECDTGTNDIPICVYKCLPTAPANGKTVSSAMEPDRYHFGQ 154

Db 61 AVRFVNCNGYKIEGDEMHCSDDGFWGKPKVCVEISCKSPVINGSPISOKI 113

QY 155 AVRFVNCNGYKIEGDEMHCSDDGFWGKPKVCVEISCKSPVINGSPISOKI 207

RESULT 3

ID W09892 standard; Protein; 484 AA.

AC W09892; (first entry)

DT 10-JUN-1999

DE Porcine E-selectin.

KW Porcine; E-selectin; diagnosis; transplant rejection; xenotransplant;

KW cytokine; donor organ endothelium; inflammatory cell; graft tissue;

KW acute cellular allograft rejection; cell adhesion.

OS Sus scrofa.

PN US891645-A.

PD 06-APR-1999.

PR 01-JUN-1994; 252493.

PR 01-JUN-1994; US-252493.

PA (ALEX-) ALEXON PHARM INC.

PI Evans MJ, Matlis LA, Rollins S, Rother RP;

DR WPI; 99-253853/21.

DR N-PSDB; X19809.

PT Nucleic acids encoding the porcine E-selectin protein - useful in

PT diagnosing and treating rejection of xenotransplanted pig organs

PS Claim 1; Column 47-50; 30pp; English.

CC The present sequence is porcine E-selectin (E-sel). The nucleotide

CC sequence encoding E-sel may be used in the diagnosis, prevention and

CC treatment of tissue rejection following xenotransplants of pig organs

CC and also in the monitoring (by specific measurement of the amount of

CC E-sel in the blood of the patient) and modulation of immune responses to

CC the transplants. Cytokine-induced expression of E-sel by donor organ

CC endothelium, contributes to the binding and subsequent transmigration of

CC inflammatory cells into the graft tissue that occurs during acute

CC cellular allograft rejection. Therefore, preventing expression of E-sel

CC on the donor tissue may reduce the risk of rejection by modulating E-sel

CC mediated cell adhesion. Fragments of the nucleotide sequence encoding

CC E-sel may be used as polymerase chain reaction (PCR) primers or

CC hybridization probes for the identification and/or isolation of the

CC porcine E-sel gene from genomic DNA. Antisense sequences may be used to

CC inhibit the expression of E-sel. DNA constructs may be used to introduce

CC antisense sequences to a cell, or to inactivate the E-sel gene (e.g.

CC knock-out constructs). Cells in which the expression of E-sel has been

CC prevented may be used as disease models or to produce transgenic animals

CC (e.g. pigs) from which organs may be harvested for transplantation with

CC a reduced risk of rejection.

SQ Sequence 484 AA;

Query Match Best Local Similarity 19.6%; Score 308; DB 1; Length 484;

Matches 52; Conservative 38; Mismatches 82; Indels 8; Gaps 6;

Db 196 QSLFNTTCAFECKEPELIGPEHLIOCTSSGWDGKKP--TKAVICDYGHPQNDVSC 253

QY 22 QTEPGTQAIYKCRPGRSGNVIMVC-RKGEWVALPLKCKRPGGDPFPTFTU 80

Db 254 NHSSIGFAVYKSTHFCAGGFGIQAQI-ECTAAGQWTOAPVCAVYKCPAVSOPKNG 312

QY 81 TGVNVFEYGAAY--YTCNNGYQLGELINRECDTGT-NTNDIPICVYKCLPTAPENG 137

Db 313 -LVFTHSPTEGFEYKSCAFSEEGPELRSNAOLCTSGOGWTOVPCVQVQCSLEY 371

QY 138 KIVSAMEPREHFGAVAFVNCNGYKIEGDEMHCSDDGFWGKPKVCVEISCKSPV 197

RESULT 4

ID R65216 standard; Protein; 830 AA.

AC R65216; (first entry)

DT 04-OCT-1995

DE P-selectin.

KW Control elements; gene expression; endothelial cells;

KW megakaryocytes; rheumatoid arthritis; ischaemic injury;

KW atherosclerosis; bacterial sepsis; tumour metastases.

OS Homo sapiens.

PN W09506118-A.

PD 02-MAR-1995.

PF 19-AUG-1994; U09395.

PR 20-AUG-1993; US-110158.

PA (OKLA) UNIV OKLAHOMA STATE.

PI McEver RP, Pan J;

DR WPI; 95-106847/14.

PT Expression control elements in the 5'-flanking region of the

PT P-selectin gene - for regulating gene expression in endothelial

PT cells and mega:karyocyte(s)

PS Disclosure; Fig 1; 88pp; English.

CC The sequence is that of the P-selectin gene product from endothelial

CC cells, determined from a composite of four overlapping cDNAs: lambda

CC GMP1-lambda GMP4. The P-selectin gene or its fragments may be used

CC to inhibit P-selectin expression, thus controlling inflammatory and

CC haemostatic processes in e.g. rheumatoid arthritis, ischaemic injury,

CC bacterial sepsis, tumour metastases, atherosclerosis, etc.

SQ Sequence 830 AA;

Query Match Best Local Similarity 19.6%; Score 308; DB 1; Length 830;

Matches 55; Conservative 44; Mismatches 94; Indels 12; Gaps 12;

Db 199 ECGEELPOHV-LIMCSPRLNFSNCSCHCTDGIQVNGPSKLEIASGIWTN-RP-P 255

QY 2 DCENLPPRRNRIELTGSVSDGTPEGTQAIYKCRPGRSGNVIMVC-RKGEWVALPLR 60

Db 256 QCLAQC-PPKIDBERNMICLSAKAFHOSSCSFCEGFEALVVG-EVVOGASGVWT 313

QY 61 KQKRPCCGHPDTP-FGFT-LTGVNVFEYGAAYVITCNGYQLGELINRECDTGT-WT 117

Db 314 APAVKACVOCQHLEAPSECTM-DC-VHPLTARVSGSCKEFCQPGYRVLGMLMCLD 371
 118 NDIPICEVAVCLPVTAPENCKIVSSAMEPDREYHFGQAVRFVCSGKIEGDEMHCSDD 177
 Db 372 GHWASAPLPTCEATSCPELSPVHGS 396
 QY 178 GWSKSKPKVEIISCKSPDV-INS 201

RESULT 5
 ID R08116 standard: protein; 610 AA.

AC R08116: 27-FEB-1991 (first entry)
 DE Endothelial cell-leukocyte adhesion molecule 1 from PCDM8 clone 6,
 DE PSQ148 and PSQ149.
 KM Endothelial cell-leukocyte adhesion molecule 1; ELAM1; antibodies;
 KM molecule involved in leucocyte adhesion; MILA; inflammation.
 OS Homo sapiens.
 PN M09013300-A.
 PD 15-NOV-1990.
 PE 27-APR-1990; U02357.
 PR 28-APR-1989; US-345151.
 PR 01-JUN-1989; US-359516.
 PR 18-DEC-1989; US-452675.
 PA (BIOC-) BIOGEN INC.
 PI Hession C, Lobb RR, Goetz SE, Born L, Benjamin CD.
 PI Rosa MD.
 DR WPI; 90-361248/48.
 DR N-PSDB; Q06686.
 PT Endothelial cell adhesion mol. - MILAs and DNA encoding them and
 PT inhibition-detection of binding of leukocytes to endothelial
 PT cells.
 PS Disclosure: Fig 1 (A-E); 136pp; English.
 PS The protein possesses a hydrophobic N-terminal sequence
 CC characteristic of a signal sequence. It is predicted that the mature
 CC N-terminal amino acid will be tryptophan at position 22. The
 CC extracellular domain of the polypeptide is ca. 554 amino acids including
 CC the signal sequence and is followed by a hydrophobic transmembrane
 CC region of 24 amino acids. The protein possesses a short, charged
 CC cytoplasmic tail of 32 amino acids. It is noted that the protein
 CC is cysteine-rich and contains eleven N-glycosylation sites.
 CC ELAM can be used to identify molecules which inhibit binding of
 CC leucocytes to cells, and such binding can be inhibited
 CC (i.e. inflammation can be treated) using an ELAM (or fragment),
 CC antibodies which recognise MILA, ELAM ligands or their fragments,
 CC carbohydrates which bind to ELAM and antibodies which recognise ELAM.
 CC See also Q06686-91.
 SQ Sequence 610 AA:

Query Match 18.9%; Score 298; DB 1; Length 610;
 Best Local Similarity 27.4%; Pred. No. 1,43e-20;
 Matches 45; Conservative 33; Mismatches 79; Indels 7; Gaps 6;

Db 210 CDGGLPSSMETWQCSSEWSA--PIPAQNVVEDCAVTNPANGVEFCQNGSPFWMTT 267
 QY 34 CRGYSLSGLVNIWCKR-GEWVALNPLRKQKRCRPGHPDPTFGFTT-LTGNNVEYGVK 91
 Db 268 CTFDCEGEFLMAQSL-OCTSSGMNDNEKPPCKAVTCANVQPOPGS-VRCSSHPAGFF 325
 QY 92 AVYTCEGQQLGEINVRCDTIG-WTNDIPICEVVKCLPTVAPENCKIVSSAMEPDREY 150
 Db 326 TFFSSCNFTCEGFMLOGPAOVECTTGGOWTOOIPVCEAFQCTA 369
 QY 151 HFGQAVRFVCSGKIEGDEMHCSDDGFWSKPKVCVEISCKS 194

RESULT 6
 ID R05494 standard: protein; 610 AA.
 AC R05494:
 DT 22-OCT-1990 (first entry)
 DE Endothelial leukocyte adhesion molecule-1 (ELAM-1).
 DE Endothelial leukocyte adhesion molecule; ELAM-1; inflammation;
 KW

KW Leukaemia; lymphoma; metastasis; cancer; ds.
 OS Homo sapiens.
 PN M09005786-A.
 PD 31-MAY-1990.
 PE 14-NOV-1989; 005073.
 PR 14-NOV-1988; US-270873.
 PA (GEOH-) Gen Hospital Corp.
 PA (BRIG-) Brigham Women's Hosp.
 PI Bevilacqua MP, Grimprone MA, Seed B, Stengelin S;
 DR WPI; 90-193436/25.
 DR N-PSDB; Q04910.
 PT Cloned genes expressing endothelial-leukocyte adhesion molecule-1 -
 PT useful in treating inflammation, post reperfusion injury,
 PT bacterial infection, vasculitis, leukaemia etc.
 PS Claim 13; 36; 54pp; English.
 CC ELAM-1 is useful in treatment of inflammation, post reperfusion
 CC injury, vasculitis etc. It may also be useful in treatment and
 CC diagnosis of leukaemia, lymphomas and metastatic spread of tumour
 CC cells, as well as in mediation of cell-cell interactions in
 CC embryogenesis and organ development, in neoplasia, and wound and
 CC tissue regeneration.
 SQ Sequence 610 AA:

Query Match 18.9%; Score 298; DB 1; Length 610;
 Best Local Similarity 27.4%; Pred. No. 1,43e-20;
 Matches 45; Conservative 33; Mismatches 79; Indels 7; Gaps 6;

Db 210 CDGGLPSSMETWQCSSEWSA--PIPAQNVVEDCAVTNPANGVEFCQNGSPFWMTT 267
 QY 34 CRGYSLSGLVNIWCKR-GEWVALNPLRKQKRCRPGHPDPTFGFTT-LTGNNVEYGVK 91
 Db 268 CTFDCEGEFLMAQSL-OCTSSGMNDNEKPPCKAVTCANVQPOPGS-VRCSSHPAGFF 325
 QY 92 AVYTCEGQQLGEINVRCDTIG-WTNDIPICEVVKCLPTVAPENCKIVSSAMEPDREY 150
 Db 326 TFFSSCNFTCEGFMLOGPAOVECTTGGOWTOOIPVCEAFQCTA 369
 QY 151 HFGQAVRFVCSGKIEGDEMHCSDDGFWSKPKVCVEISCKS 194

RESULT 7
 ID R05538 standard: protein; 610 AA.
 AC R05538:
 DT 24-OCT-1990 (first entry)
 DE Endothelial-leukocyte adhesion molecule (ELAM-1).
 KW Endothelial-leukocyte adhesion molecules; ELAM-1; cytotoxin;
 KW cytokine; cancer; leukemia; ds.
 OS Homo sapiens.
 PN M09005539-A.
 PD 31-MAY-1990.
 PE 14-NOV-1989; 005075.
 PR 14-NOV-1988; US-270860.
 PA (BRIG-) Brigham & Women's Ho.
 PI Bevilacqua MP, Grimprone MA;
 DR WPI; 90-193270/25.
 DR N-PSDB; Q04776.
 PT Antibodies against ELAM-1 to treat inflammation, infection,
 PT and for detecting expression of this endothelial leukocyte
 PT adhesion protein.
 PS Disclosure: 44pp; English.
 CC ELAM specific antibodies, attached to a suitable drug eg. a
 CC cytotoxin can direct the action of the drug to sites of
 CC inflammation, tissue damage and infection. The same Abs can be
 CC used to treat leukaemia, lymphoma, vasculitis and metastatic spread
 CC of cancer cells. Labelled ELAM specific Abs can be used to detect
 CC expression of the adhesion molecules.
 SQ Sequence 610 AA:

Query Match 18.9%; Score 298; DB 1; Length 610;
 Best Local Similarity 27.4%; Pred. No. 1,43e-20;
 Matches 45; Conservative 33; Mismatches 79; Indels 7; Gaps 6;

Db 210 CDRGYLPSSMETNOCSSGEMSA--PIPACNVECDVATNPANGFVECFQNPSPFWNTT 267
 QY 34 CRFGYRLSNVIMVCRK-GEWVALNPLRKQCRPGHDPDPTFGFTT-LTGNNVFEYGVK 91
 Db 268 CTFDEEGFELMGAQSL-QCTSSGNMNDNEKPTCKAVTCRAVROPONGS-VRCSSHSPAGEF 325
 QY 92 AYTTCNEGQYQLLGEINVRCDTIDG-WTNDIPICEVVKCLPVTAPENKIVSSAMEPDREY 150
 Db 326 TFKSSCNFTCEBGFMLQCPAYVECTTQGTQOIPVCEAFQCTA 369
 QY 151 HFGQAVRFVNCNSGYKIEGDEEMHCSDGFWSKERKPCVEISCKS 194

RESULT 8
 ID W18839 standard; Protein: 610 AA.
 AC W18839;
 DT 05-JAN-1998 (first entry)
 DE E-selectin.
 KW murine monoclonal antibody; H4/18; H18/7; immunoglobulin; activated endothelium; treat; inflammation; microbial infection; post-reperfusion injury; malignant tumour; vasculitis; vascular smooth-muscle cell proliferative disorder; acute; chronic allograft rejection.
 OS Homo sapiens.
 PN US5632991-A.
 PD 27-MAY-1997.
 PF 14-NOV-1988; 270860.
 PR 29-DEC-1994; US-365470.
 PR 14-NOV-1988; US-270860.
 PR 13-MAR-1992; US-850802.
 PR 05-AUG-1993; US-102510.
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL.
 PI Gimbome MA;
 DR WPI: 97-297321/27.
 DR N-PSDB: 170493.
 PT Immunoconjugates based on monoclonal antibody specific for E-selectin - for targeting therapeutic agent to activated endothelium
 PS Example 2: Column 35-38; 36pp: English.
 CC This sequence is E-selectin (formerly known as endothelial leukocyte adhesion molecule-1; ELAM-1). E-selectin is identified by specific binding of murine monoclonal antibodies H4/18 and H18/7 developed against stimulated endothelial cells. New immunoconjugates of the invention comprise monoclonal antibody H18/7, or a fragment which binds to E-selectin conjugated to a therapeutic agent. The immunoconjugate is used to selectively target the therapeutic agent to activated endothelium expressing E-selectin, especially for treating inflammation, microbial infections, post-reperfusion injury, malignant tumours, vasculitis, vascular smooth-muscle cell proliferative disorders or acute or chronic allograft rejection.
 SQ Sequence 610 AA;

Query Match 18.9%; Score 298; DB 1; Length 610;
 Best Local Similarity 27.4%; Pred. No. 1,43e-20;
 Matches 45; Conservative 33; Mismatches 79; Indels 7; Gaps 6;
 Db 210 CDRGYLPSSMETNOCSSGEMSA--PIPACNVECDVATNPANGFVECFQNPSPFWNTT 267
 QY 34 CRFGYRLSNVIMVCRK-GEWVALNPLRKQCRPGHDPDPTFGFTT-LTGNNVFEYGVK 91
 Db 268 CTFDEEGFELMGAQSL-QCTSSGNMNDNEKPTCKAVTCRAVROPONGS-VRCSSHSPAGEF 325
 QY 92 AYTTCNEGQYQLLGEINVRCDTIDG-WTNDIPICEVVKCLPVTAPENKIVSSAMEPDREY 150
 Db 326 TFKSSCNFTCEBGFMLQCPAYVECTTQGTQOIPVCEAFQCTA 369
 QY 151 HFGQAVRFVNCNSGYKIEGDEEMHCSDGFWSKERKPCVEISCKS 194

DT 02-JUN-1998 (first entry)
 DE Amino acid sequence of endothelial leukocyte adhesion molecule-1.
 KW Endothelial leukocyte adhesion molecule-1; ELAM-1; inducible; cytokine; overexpression; cellular adhesion molecule; intracellular adhesion molecule-1; ICAM-1; antisense molecule; viral hepatitis; Alzheimer's disease; multiple sclerosis; immunoglobulin; cholangitis; cardiac allograft rejection.
 OS Homo sapiens.
 PN W09746671-AL.
 PD 11-DEC-1997.
 PF 22-MAY-1997; CA0347.
 PR 30-MAY-1996; US-657753.
 PA (UYBR-) UNIV BRITISH COLUMBIA.
 PI Hope MJ, Klimuk SK, Scherrer P, Sample SC;
 DR WPI: 98-042180/04.
 DR N-PSDB: V16208.
 PT Composition for treatment of conditions associated with overexpression of ICAM-1 - used to treat e.g. Alzheimer's disease, glomerulonephritis, rheumatoid arthritis etc.
 PS Disclosure: Pages 44-46; 81pp: English.
 CC The present sequence represents human endothelial leukocyte adhesion molecule-1 (ELAM-1). ELAM-1 is a membrane glycoprotein which is induced by a number of cytokines. Pathological conditions associated with the overexpression of cellular adhesion molecules, such as ELAM-1 or intracellular adhesion molecule-1 (ICAM-1), can be treated with the novel pharmaceutical composition of the invention. The composition comprises an effective amount of an ICAM-1 antisense molecule encapsulated in a lipid mixture, the lipid mixture comprising at least two members selected from phospholipids, steroids and cationic lipids. The composition is used in a method to treat CC pathological conditions associated with overexpression of ICAM-1, such as Alzheimer's disease, multiple sclerosis, viral hepatitis, cholangitis, cardiac allograft rejection, etc.
 SQ Sequence 610 AA;

Query Match 18.9%; Score 298; DB 1; Length 610;
 Best Local Similarity 27.4%; Pred. No. 1,43e-20;
 Matches 45; Conservative 33; Mismatches 79; Indels 7; Gaps 6;
 Db 210 CDRGYLPSSMETNOCSSGEMSA--PIPACNVECDVATNPANGFVECFQNPSPFWNTT 267
 QY 34 CRFGYRLSNVIMVCRK-GEWVALNPLRKQCRPGHDPDPTFGFTT-LTGNNVFEYGVK 91
 Db 268 CTFDEEGFELMGAQSL-QCTSSGNMNDNEKPTCKAVTCRAVROPONGS-VRCSSHSPAGEF 325
 QY 92 AYTTCNEGQYQLLGEINVRCDTIDG-WTNDIPICEVVKCLPVTAPENKIVSSAMEPDREY 150
 Db 326 TFKSSCNFTCEBGFMLQCPAYVECTTQGTQOIPVCEAFQCTA 369
 QY 151 HFGQAVRFVNCNSGYKIEGDEEMHCSDGFWSKERKPCVEISCKS 194

RESULT 10
 ID W06881 standard; Protein: 299 AA.
 AC W06881;
 DT 18-MAR-1997 (first entry)
 DE Decay accelerating factor.
 KW Complement inhibitor; membrane co-factor protein; MCP; decay accelerating factor; DAF; chimeric protein; glycosaminoglycan; heparin; cell lysis; sepsis; adult respiratory distress syndrome; reperfusion injury; cell damage.
 OS Homo sapiens.
 PN W09634965-A2.
 PD 07-NOV-1996.
 PF 03-MAY-1996; U06301.
 PR 05-MAY-1995; US-435149.
 PA (CHIR) CHIRON CORP.
 PI Creasey AA, Innis MA, Zaror I;
 DR WPI: 96-506167/50.
 PT Chimeric proteins for inhibiting complement-mediated cell lysis - comprise membrane co-factor protein and decay accelerating factor peptide sequences
 PS Disclosure: Page 25; 33pp: English.

OY 116 WTNDIPICEVVKCLPVTAPENGRKIVSSAMEPDRHYHFGQAVRFVCSNGYRIEEDDEMHCS 175
 DB 269 VNNDGEMSGPPEPC 283
 OY 176 ---DDGFWSKEKPKC 187

RESULT 13
 ID P70048 standard; protein; 381 AA.
 AC P70048:
 DT 03-FEB-1991 (first entry)
 DE Human decay acceleration factor variant #1.
 KM Decay acceleration factor.
 FH Key Location/Qualifiers
 FT region 331..347
 PT /label-putative transmembrane region
 PD EP-244267-A.
 PN 04-NOV-1987.
 PR 01-MAY-1987; 303944.
 PA 02-MAY-1986; US-859107.
 PI (GERTH) Genentech Inc.
 DR Caras IW;
 DR WPI: 87-308481/44.
 DR N-PSDB: N70047.
 PT New decay accelerating factor variants - obtained with the factor by
 using recombinant DNA procedures.
 PS Disclosure; Page 15-17; 20pp; English.
 CC The probable phosphatidylinositol derivatisation site is Cys(330).
 CC The DAF variant is useful for treating paroxysmal nocturnal
 CC haemoglobinuria, or inflammatory or cell lytic autoimmune
 CC diseases. It may be used to ameliorate allograft rejection
 CC or autoimmune diseases. See also N70046, N70048.
 SQ Sequence 381 AA;

Query Match 18.6%; Score 293; DB 1; Length 381;
 Best Local Similarity 29.7%; Pred. No. 4,96e-20;
 Matches 58; Conservative 43; Mismatches 75; Indels 19; Gaps 14;

DB 98 C-EVPTRLNSASLKQPTITQNYFPVGTVEYECRPGYRREPSLSFKTLCLQNLKMWSTAVE 156
 OY 3 CNELPRRNTTEILTGSMSDQTY-PEGTQAIYKCRPGYSLGNVI--MYC-RKGEWVALNP 58
 DB 157 F--CKKSCPNNGEIRNGQIDVPG-IL-FGATISFSCNTGKTLGSSFCLLISGSSVQ 212
 OY 59 LKCKQKRCGHPGDPFGFTLTGNGVFEYGVKAVYTCNEGTLGELN-YR-ECDTDG- 115
 DB 213 WSDPLPECREIYCAPPQIDNG-IIQG--ERDH-YGYROSVTYACNKGFTMIGESHIYCT 268
 OY 116 WTNDIPICEVVKCLPVTAPENGRKIVSSAMEPDRHYHFGQAVRFVCSNGYRIEEDDEMHCS 175
 DB 269 VNNDGEMSGPPEPC 283
 OY 176 ---DDGFWSKEKPKC 187

RESULT 14
 ID P94773 standard; protein; 381 AA.
 AC P94773:
 DT 04-JUL-1990 (first entry)
 DE Decay accelerating factor (DAF) of clones lambda 33 and lambda 47.
 KM DAF; allograft rejection; affinity purification;
 KW autoimmune disease; ds.
 OS Synthetic.
 PN W08901041-A.
 PD 09-FEB-1989.
 PR 3-AUG-1988; 02648.
 PR 6-AUG-1987; US-083757.
 PA (GERTH) Genentech Inc.
 PI Caras I;
 DR WPI: 89-061177/08.
 DR N-PSDB: N91043.
 PT Fusion polypeptide for targeting protein to cell membrane -

PT compisesphospholipid anchor domain with heterologous
 PT polypeptide.
 PS Disclosure; 61pp; English.
 CC Recombinant DAF's are useful in treatment of inflammatory or cell lytic
 CC autoimmune diseases and allograft rejection. Useful in diagnostic
 CC compositions or in affinity purification.
 SQ Sequence 381 AA;

Query Match 18.6%; Score 293; DB 1; Length 381;
 Best Local Similarity 29.7%; Pred. No. 4,96e-20;
 Matches 58; Conservative 43; Mismatches 75; Indels 19; Gaps 14;

DB 98 C-EVPTRLNSASLKQPTITQNYFPVGTVEYECRPGYRREPSLSFKTLCLQNLKMWSTAVE 156
 OY 3 CNELPRRNTTEILTGSMSDQTY-PEGTQAIYKCRPGYSLGNVI--MYC-RKGEWVALNP 58
 DB 157 F--CKKSCPNNGEIRNGQIDVPG-IL-FGATISFSCNTGKTLGSSFCLLISGSSVQ 212
 OY 59 LKCKQKRCGHPGDPFGFTLTGNGVFEYGVKAVYTCNEGTLGELN-YR-ECDTDG- 115
 DB 213 WSDPLPECREIYCAPPQIDNG-IIQG--ERDH-YGYROSVTYACNKGFTMIGESHIYCT 268
 OY 116 WTNDIPICEVVKCLPVTAPENGRKIVSSAMEPDRHYHFGQAVRFVCSNGYRIEEDDEMHCS 175
 DB 269 VNNDGEMSGPPEPC 283
 OY 176 ---DDGFWSKEKPKC 187

RESULT 15
 ID W73505 standard; protein; 381 AA.
 AC W73505:
 DT 01-MAR-1999 (first entry)
 DE Decay accelerating factor protein.
 KM DAF; decay accelerating factor; human; complement protein; gene therapy;
 KW viral vector; ds.
 OS Homo sapiens.
 PN J10313865-A.
 PD 02-DEC-1997; 125965.
 PR 15-MAY-1997; JP-125965.
 PA (DINA-) DINABEKRU KENKYUSHO KK.
 DR WPI: 99-074147/07.
 DR N-PSDB: V08935.
 PT Vector having complement controlling factor - useful for gene
 PT therapy
 PS Example 3; Page 8-9; 15pp; Japanese.
 CC This sequence is the human decay accelerating factor (DAF) protein.
 CC DAF can be used in the viral vector of the invention. The viral vector
 CC contains a factor controlling the function of a human complement protein,
 CC particularly a membrane combining type protein. The viral vector, which
 CC is stable in situ, is useful for gene therapy.
 SQ Sequence 381 AA;

Query Match 18.6%; Score 293; DB 1; Length 381;
 Best Local Similarity 29.7%; Pred. No. 4,96e-20;
 Matches 58; Conservative 43; Mismatches 75; Indels 19; Gaps 14;

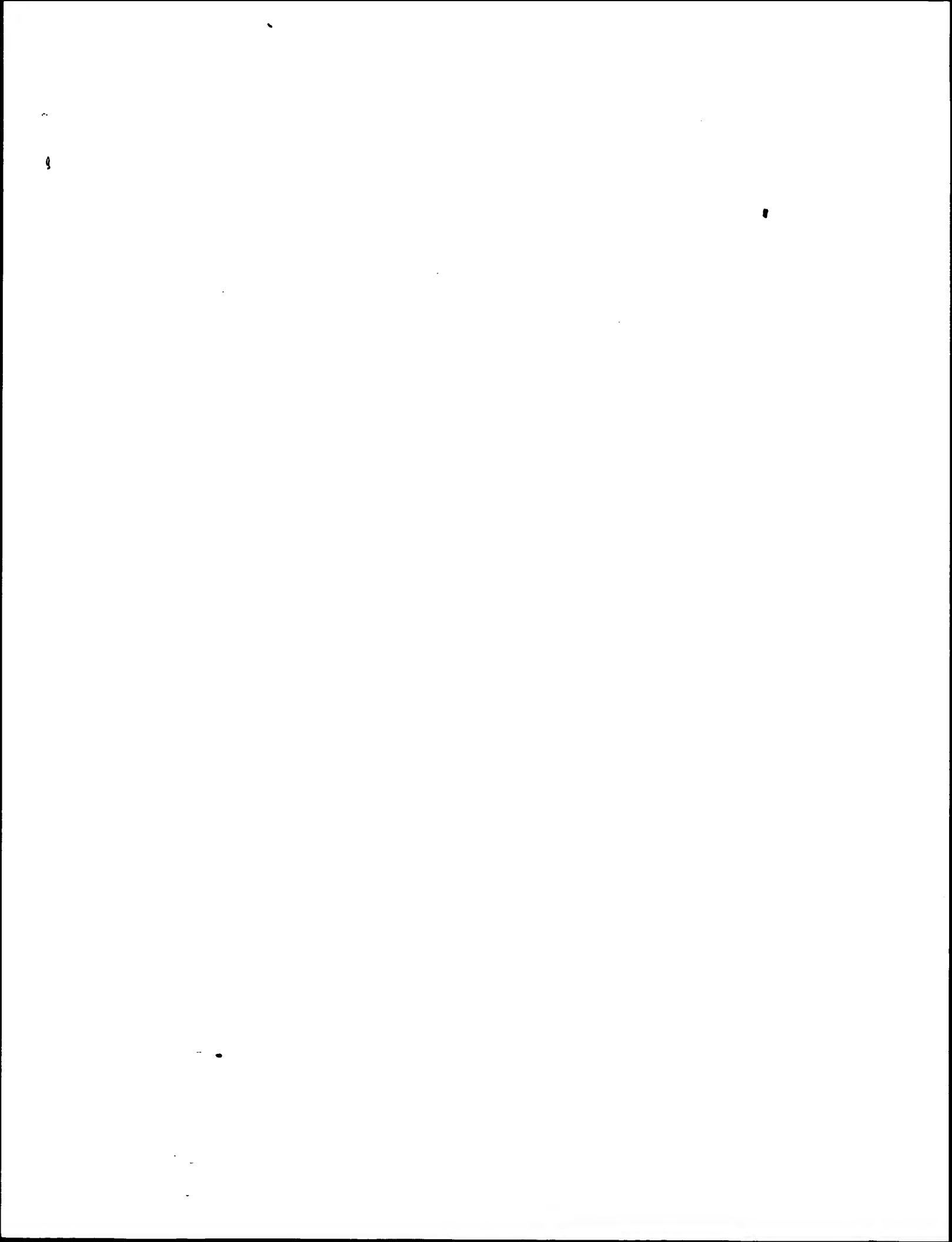
DB 98 C-EVPTRLNSASLKQPTITQNYFPVGTVEYECRPGYRREPSLSFKTLCLQNLKMWSTAVE 156
 OY 3 CNELPRRNTTEILTGSMSDQTY-PEGTQAIYKCRPGYSLGNVI--MYC-RKGEWVALNP 58
 DB 157 F--CKKSCPNNGEIRNGQIDVPG-IL-FGATISFSCNTGKTLGSSFCLLISGSSVQ 212
 OY 59 LKCKQKRCGHPGDPFGFTLTGNGVFEYGVKAVYTCNEGTLGELN-YR-ECDTDG- 115
 DB 213 WSDPLPECREIYCAPPQIDNG-IIQG--ERDH-YGYROSVTYACNKGFTMIGESHIYCT 268
 OY 116 WTNDIPICEVVKCLPVTAPENGRKIVSSAMEPDRHYHFGQAVRFVCSNGYRIEEDDEMHCS 175
 DB 269 VNNDGEMSGPPEPC 283
 OY 176 ---DDGFWSKEKPKC 187

Fri Jun 9 10:53:44 2000

US-09-316-163-9.rag

Page 7

Search completed: Thu Jun 8 21:37:19 2000
Job time : 11 secs.



 W39154
 (TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Merch_p protein - protein database search, using Smith-Waterman algorithm
 Run on: Thu Jun 8 21:40:50 2000; Maspar time 9.79 Seconds
 641.132 Million cell updates/sec
 Tabular output not generated.

Title: >US-09-316-163-10
 Description: (1-265) from US09316163.pep
 Perfect Score: 2029
 Sequence: 1 EDCEMLPPRRMTTEILTSWS.....EKSCDNPYIPNGDYSPLRIK 265

Scoring table: PAM 150
 Gap 11

Searched: 188963 seqs/23686106 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: a:geneseq36
 1:geneseqp

Statistics: Mean 32.240; Variance 119.188; scale 0.271

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1429	70.4	240	1	Human partial Complement	1.09e-147
2	1298	64.0	216	1	Clone PRB9FH410 CFH r	1.28e-132
3	382	18.8	581	1	Human C4 binding prote	4.34e-29
4	371	18.3	263	1	Deduced sequence of co	6.80e-28
5	344	17.0	830	1	P-selectin.	5.67e-25
6	341	16.8	302	1	Herpesvirus secreted g	1.19e-24
7	341	16.8	302	1	Herpesvirus salm1r1 SC	1.19e-24
8	341	16.8	360	1	Herpesvirus membrane g	1.19e-24
9	341	16.8	360	1	Herpesvirus salm1r1 mc	1.19e-24
10	335	16.5	376	1	CD46 wild-type.	5.29e-24
11	335	16.5	377	1	CD46 construct subSC3	5.29e-24
12	335	16.5	377	1	Human MCP.	5.29e-24
13	335	16.5	384	1	Human membrane cofacto	5.29e-24
14	335	16.5	384	1	Human CD46.	5.29e-24
15	333	16.4	279	1	rSMCP protein SEQ ID N	8.68e-24
16	333	16.4	279	1	Membrane cofactor prot	8.68e-24
17	332	16.4	324	1	CD46 from clone pm5.8.	1.11e-23
18	332	16.4	324	1	CD46 from clone pm5.1.	1.11e-23
19	332	16.4	377	1	CD46 from clone pm5.1.	1.11e-23
20	330	16.3	254	1	Membrane co-factor pro	1.82e-23
21	330	16.3	254	1	Membrane co-factor pro	1.82e-23
22	330	16.3	357	1	Human membrane cofacto	1.82e-23
23	330	16.3	577	1	Membrane co-factor pro	1.82e-23

RESULT ID	Score	Query Match	Length	ID	Description	Pred. No.
1	1429	70.4	240	1	Human partial Complement	1.09e-147
2	1298	64.0	216	1	Clone PRB9FH410 CFH r	1.28e-132
3	382	18.8	581	1	Human C4 binding prote	4.34e-29
4	371	18.3	263	1	Deduced sequence of co	6.80e-28
5	344	17.0	830	1	P-selectin.	5.67e-25
6	341	16.8	302	1	Herpesvirus secreted g	1.19e-24
7	341	16.8	302	1	Herpesvirus salm1r1 SC	1.19e-24
8	341	16.8	360	1	Herpesvirus membrane g	1.19e-24
9	341	16.8	360	1	Herpesvirus salm1r1 mc	1.19e-24
10	335	16.5	376	1	CD46 wild-type.	5.29e-24
11	335	16.5	377	1	CD46 construct subSC3	5.29e-24
12	335	16.5	377	1	Human MCP.	5.29e-24
13	335	16.5	384	1	Human membrane cofacto	5.29e-24
14	335	16.5	384	1	Human CD46.	5.29e-24
15	333	16.4	279	1	rSMCP protein SEQ ID N	8.68e-24
16	333	16.4	279	1	Membrane cofactor prot	8.68e-24
17	332	16.4	324	1	CD46 from clone pm5.8.	1.11e-23
18	332	16.4	324	1	CD46 from clone pm5.1.	1.11e-23
19	332	16.4	377	1	CD46 from clone pm5.1.	1.11e-23
20	330	16.3	254	1	Membrane co-factor pro	1.82e-23
21	330	16.3	254	1	Membrane co-factor pro	1.82e-23
22	330	16.3	357	1	Human membrane cofacto	1.82e-23
23	330	16.3	577	1	Membrane co-factor pro	1.82e-23

ALIGNMENTS

RESULT 1
 ID W39154 standard: Protein; 240 AA.
 AC W39154;
 DT 27-APR-1998 (first entry)
 DE Human partial Complement factor H protein fragment 1.
 KW Complement factor H; tumour associated antigen; renal cancer;
 KW urogenital cancer; medicament; modulator.
 OS Homo sapiens.
 PN WO9738136-A1.
 PD 16-OCR-1997.
 PE 09-APR-1997; U05710.
 PR 06-MAR-1997; US-812481.
 PR 09-APR-1996; US-015083.
 PR 09-APR-1996; US-630048.
 PR 06-MAR-1997; US-038614.
 PA (BARD-) BARD DIAGNOSTIC SCI INC.
 PI Enfield DL, Hass GM, Kinders RJ;
 PI WPI: 97-512742/47.
 DR N-PSDB: V02790.
 PT Treating or screening for cancer, e.g. renal or urogenital cancer
 PT by modulating or detecting tumour associated human complement factor
 PT H related antigen, or nucleic acid encoding it
 PS Example 6B: F1d 5B: 104pp; English.
 CC This partial protein sequence represents a region of the human
 CC tumour-associated complement factor H (CFH). This sequence is used
 CC in the identification of complement factor H related proteins and
 CC such proteins isolated from clone PRB9FH410 (see W39154). The detection of
 CC treatment of renal or urogenital cancer, e.g. bladder, cervical or
 CC prostate cancer. Agents that may modulate this antigen could be used in
 CC the manufacture of a medicament for the treatment of a tumour cell.
 SQ Sequence 240 AA:
 Query Match 70.4%; Score 1429; DB 1; Length 240;
 Best Local Similarity 100.0%; Pred. No. 1.09e-147;
 Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 FTLLGVNFEYGVAVYTCNEGVOYLLGEINVRCDIDGWTNDIPICEVVKCLPVTAPENG 60
 78 FTLLGVNFEYGVAVYTCNEGVOYLLGEINVRCDIDGWTNDIPICEVVKCLPVTAPENG 137
 Oy 61 KIVSSAMEPPREYHFGAVRVCNSGYKIGDEBHCSDGCFNSKPKVETISCSPPV 120
 138 KIVSSAMEPPREYHFGAVRVCNSGYKIGDEBHCSDGCFNSKPKVETISCSPPV 197
 Db 121 INGSISGKITYKNEPFOYKCNMGYSESGDGVAVCVESGMRPLPSCSEKSCNPPIPN 180

QY 198 INGSPIKIKYKENERFOYKCMNGYERSDGAVCTESGMRPLPSCSEKSCDNPIYING 257
 Db 181 DYSPLRIK 188
 QY 258 DYSPLRIK 265

RESULT 2
 ID W39155 standard; Protein; 216 AA.

AC W39155;
 DT 27-APR-1998 (first entry)
 DE Clone pRRB9FH410 CFH related protein fragment.
 KW Complement factor H; tumour associated antigen; renal cancer;
 KW urogenital cancer; medicament; modulator.
 OS Synthetic.
 PN WO9738136-A1.
 PD 16-OCT-1997.
 PF 09-APR-1997; U057110.
 PR 06-MAR-1997; US-812481.
 PR 09-APR-1997; US-015083.
 PR 09-APR-1996; US-630048.
 PR 06-MAR-1997; US-038614.
 PA (BARD-) BARD DIAGNOSTIC SCI INC.
 PI Enfield DL, Hass GM, Kinders RJ;
 DR WPI: 97-512742/47.
 DR N-PSDB: V02791.

PT Treating or screening for cancer, e.g. renal or urogenital cancer -
 PT by modulating or detecting tumour associated human complement factor -
 PT H related antigen, or nucleic acid encoding it
 PS Example 6B; Fig 6B; 104pp; English.
 CC This partial protein is found in clone pRRB9FH410 and represents a
 CC complement factor H related protein with homology to a region of the
 CC human tumour-associated complement factor H (CFH). The detection of this
 CC protein and a CFH antigen can be used in screening or for the treatment
 CC of renal or urogenital cancer, e.g. bladder, cervical or prostate cancer.
 CC Agents that may modulate this antigen could be used in the manufacture of
 CC a medicament for the treatment of a tumour cell.
 SQ Sequence 216 AA;

Query Match 64.0%; Score 1298; DB 1; Length 216;
 Best Local Similarity 99.4%; Pred. No. 1.28e-132;
 Matches 170; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 TCNEGYOLLEINRECDIDMTNDIPICFVVKCLPYTAPENKITYSSAMEPDRHFQ 60
 QY 95 TCNEGYOLLEINRECDIDMTNDIPICFVVKCLPYTAPENKITYSSAMEPDRHFQ 154
 Db 61 AVRFVCSGKIRIGDEMGHSDGFWGKEKPKCVEISCKSPDYINGSPISOKIYKENER 120
 QY 155 AVRFVCSGKIRIGDEMGHSDGFWGKEKPKCVEISCKSPDYINGSPISOKIYKENER 214
 Db 121 FOYKCMNGYERSDGAVCTESGMRPLPSCSEKSCDNPIYINGDYSPLRIK 171
 QY 215 FOYKCMNGYERSDGAVCTESGMRPLPSCSEKSCDNPIYINGDYSPLRIK 265

RESULT 3
 ID R13490 standard; Protein; 581 AA.
 AC R13490;
 DT 30-OCT-1991 (first entry)
 DE Human C4 binding protein.
 KW C4bp: monomer; complement protein; pIOD.C4bp.3; SCR;
 KW short consensus repeat.
 OS Homo sapiens.
 FH Key
 FH peptide
 FT 1..32
 FT /label= signal_peptide
 FT 33..581
 FT /label= C4bp
 FT 33..93
 FT /label= SCR8
 FT 94..155
 FT /label= SCR7

FT region
 FT 156..219
 FT /label= SCR6
 FT 220..279
 FT /label= SCR5
 FT 280..345
 FT /label= SCR4
 FT 346..406
 FT /label= SCR3
 FT 407..464
 FT /label= SCR2
 FT 465..523
 FT /label= SCR1
 FT 524..581
 FT /label= C4bp_core
 FT /note= "responsible for multimer assembly"
 FT disulfide_bond
 FT 34..80
 FT /note= "intradomain"
 FT disulfide_bond
 FT 65..92
 FT /note= "intradomain"
 PN MO9111461-A.
 PD 08-AUG-1991.
 PF 28-JAN-1991; U00567.
 PR 26-JAN-1990; US-470888.
 PA (BIOG-) BIOGEN INC.
 PI Pasek MP, Winkler G, Liu TR;
 DR WPI: 91-252613/34.
 DR N-PSDB: Q13242.

PT New C4 binding protein fusion proteins and DNA encoding them -
 PT comprise assemblies of C4bp monomers linked to functional moiety,
 PT e.g. A2T, useful as delivery vehicles in diagnosis and therapy
 PS Example 1; Fig 1; 105pp; English.
 CC This sequence was deduced from human hepatocyte (Hep G2) cDNA
 CC obtained following PCR amplification. The protein is a monomer
 CC containing 8 SCRs. Each SCR forms a looped domain due to the
 CC presence of two intradomain disulfide bonds (only the disulfide
 CC bonds of SCR8 are labelled in the Features Table). Within each SCR,
 CC the first cysteine residue bonds with the fourth. This secondary structure
 CC cysteine residue bonds with the fourth. This secondary structure is
 CC responsible for the conformational flexibility of the C4bp monomer.
 CC The invention covers fusion proteins in which the monomer sequence,
 CC or a specified subfragment of it, e.g. having 5', 4', 3 or 1 SCR(s)
 CC is fused to the C-terminal of a protein such as a viral receptor,
 CC cell ligand, a bacterial, viral or parasitic immunogen, enzyme,
 CC cytokine, toxin, etc. See also Q13243-51.
 SQ Sequence 581 AA;

Query Match 18.8%; Score 382; DB 1; Length 581;
 Best Local Similarity 31.3%; Pred. No. 4.34e-29;
 Matches 79; Conservative 52; Mismatches 98; Indels 22; Gaps 20;

Db 51 TETRFGTTLTKYTCPLPGYRSHSTQTLFNCSDGEMV-YNTFCIT-KR-CRHPGELRNQ 107
 QY 20 SDQTYPRGTQAIYKCRGY-RSLGNVIMCRK-GEWALNPLKCKCRPGHGDDPFPT 77
 Db 108 VEIKT-DL-SFGSIEPSCSEGFLLIGSTTSKCEVODRGVMSHPLPQCEIYKCKPPDI 165
 QY 78 FTLLGAVVEFYKAVYTCNEGYOLLEINR-EC-DTD-GWTNDIPICFVVKCLPYTAP 134
 Db 166 RNGR--HSG-E-ENFYAYGGSVYSCDPRSLIGHASISCTYENETIGWRSPPCERT 221
 QY 135 ENGRIVSSAMEPDRHFQAVFVCSGKIRIGDEMGHSDGFWGKEKPKCVEI 190
 Db 222 TCRKPDVSHGEMVSGFPIYKDTIVFCGNGFVLRGSSVYHCDADSKWNSPPACEPN 281
 QY 191 SCKSPDYINGSPISQ-KIYKENERFOYKCMNGYERSDGAVCTESGMRPLPSCSEK 247
 Db 282 SCINLPDIPA 292
 QY 248 SCIN-PYING 257

RESULT 4
 ID P92003 standard; Protein; 263 AA

Query Match	16.8%;	Score 341;	DB 1;	Length 360;
Best Local Similarity	31.9%;	Pred. No. 1.19e-24;		
Matches	74;	Mismatches 91;	Indels 22;	Gaps 17

RESULT	10
ID	R93939 standard; Protein; 376 AA
NC	000000

DT 20-MAY-1996 (first entry)
DE CD46 wild-type.
KW CD46: recombinant protein; short consensus repeat; SCR;
KW regulator of complement activation; transgenic animal; pig;
KW organ transplantation.
OS Homo sapiens.
PN W09606937-A1.
PD 07-MAR-1996.
PE 30-AUG-1995; AU0553.
PR 30-AUG-1994; AU-007724.
PA (AUST-) AUSTIN RES. INST.
PI Christiansen D, Loveland B, McKenzie IFC, Milland J;
DR WPI: 96-160368/16.
DR N-PSDB: T17595.
PT Increasing prodn. of recombinant proteins, esp. CD46 - by reducing
PT the amt. of A and/or T in an A and/or T rich region of encoding gene
PS exon
PS Disclosure: Page 32-33; 60pp; English.
CC A cDNA sequence (T17595) codes for wild-type CD46 (R93939), a

Query Match	16.5%;	Score 335;	DB 1;	Length 376;
Best Local Similarity	31.0%;	Pred. No. 5.29e-24;		
Matches	81;	Conservative	54;	Mismatches 99;
			Indels 27;	Gaps 22;

RESULT	11
ID	R93941 standard; Protein; 377 AA

DT 21-MAY-1996 (first entry)
 DE CD46 construct subSCR3.
 KW CD46, recombinant protein; short consensus repeat; SCR;
 KM regulator of complement activation; transgenic animal; pig;
 organ transplantation.
 OS Synthetic.
 PN MO9606937-A1.
 PD 07-MAR-1996.
 PR 30-AUG-1995; AU0553.
 PR 30-AUG-1994; AU-007724.
 PA (AUST-) AUSTRIAN RES INST.
 PI Christiana D. Loveland B, McKenzie IFC, Millard J;
 DR WPI: 96-160368/16.
 DR N-PSDB: T17597.
 PT Increasing prodn. of recombinant proteins, esp. CD46 - by reducing
 the amt. of A and/or T in an A and/or T rich region of encoding gene
 exon
 PS Claim 12, Page 36-37; 60pp; English.
 CC CD46 subSCR3 (R93941) is the product of a cDNA construct
 CC (T17597) chbd. by splice overlap extension PCR of wild-type CD46
 CC cDNA (T17595). The A/T content of A/T-rich exon 5 of the CD46
 CC gene, encoding the short consensus repeat 3 (SCR3) region of
 CC CD46, is lowered. This results in improved prodn. in eukaryotic
 CC host cells, e.g. CHO-K1 and COS-7. CD46 subSCR3 is used to prevent
 CC complement- or inflammation-mediated tissue damage, to improve
 CC immunity to tumours or viruses, to control fertilisation and to
 CC prevent spontaneous abortion. Expression in transgenic animals,
 CC esp. pigs, provides organs suitable for transplantation.
 Sequence 377 AA;
 SQ

	Query Match	16.58;	Score 335;	DB 1;	Length 377;	
	Best Local Similarity	31.08;	Pred. No. 5,29e-24;			
	Matches	81;	Conservative	54;	Mismatches 99;	Indels 27; Gaps 22
Db	33 DACEE-PPTFAMEMLICK-KPKYYIEIGERVDYCKCKGGYYIPPLATHTICDNRHTEFLVFS	90				
	: : :	:	:	:	:	: : :::

QY 1 EDCNEIPRRNTEILTGNSDQTYPEGTQAIKCRPGYSLGNV-I-MVC-RKGEWALN 57
 Db 91 D-DACYRETCPYIRD-PLNGQAVPANGTYEFGYOMHFIENEGYIIGELITCELKGSVA 148
 QY 58 PLKRCQKRCRPGHGDTPFGFTILGNGVFEYGVKAVYTNEGQYQILGE-INYRE-CDTGG 115
 Db 149 IMSGRPIECYKVLCTPPKIKNGKHTFSEYEV-FEY-L-DAVYSCDPAAGPDPFSLIGE 205
 QY 116 -WTNDIPICEVVKCLPVTAPENCKIYSSAMEPDRYHFGQAVFYCN-S-G---YKIEGD 169
 Db 206 STIYCGNSVMSRAAPCEKAVKCRFPVYENGKQISFGKKEFYKATVME--CDGKF-YL 262
 QY 170 EHMCHSDGDFWMSKEKRCVCEISCKSPDYINGSPIS---QKIYKENERFYQKCNMGYEYS 226
 Db 263 DGSDDTYCDNSSTWDPVPVPC 283
 QY 227 ERGDA-VC-TESGM-RPLPSC 244

RESULT 12

ID W27484 standard; Protein; 377 AA.
 AC W27484.
 DT 14-APR-1998 (first entry)
 DE Human MCP.
 KW Human MCP; surface; homologous complement restriction factor;
 HCRF; medical apparatus; medical dressing; surgical equipment;
 KW diagnostic kit; prevention; purification device; reduction;
 KW complement activation.
 OS Homo sapiens.
 PN W09735886-A1.
 PD 02-OCT-1997.
 PF 12-MAR-1997; G00684.
 PR 31-OCT-1996; GB-022694.
 PR 22-MAR-1996; GB-006073.
 PR 28-MAR-1996; GB-006516.
 PA (IMOT-) IMOTRAN LTD.
 PI Watkins NJ.
 DR WPI: 97-489571/45.
 DR N-PSDB: T90306.
 PT Surfaces which prevent or reduce complement activation - having
 PT surface bound homologous complement restriction factor
 PS Example 1; Fig 1A; 50pp; English.
 CC The present sequence is human MCP, which was used in the
 CC development of a novel surface bound to a homologous complement
 CC restriction factor (HCRF), where the surface is not a surface to
 CC which HCRF binds in vivo. The surface can be used in medical
 CC apparatus (e.g. extra-corporeal circulation systems, tubing,
 CC valves, membranes, pumps, oxygenators, catheters, cannulas, fluid
 CC reservoirs or prostheses), medical dressings, surgical equipment,
 CC diagnostic kits (e.g. kits for determining whether a patient has
 CC an abnormality which prevents the complement system from working
 CC normally or from being properly regulated) and purification
 CC devices (e.g. for purifying complement components). The surface
 CC can be used to reduce or prevent activation of complement, and to
 CC reduce morbidity due to complement activation.
 SQ Sequence 377 AA;

Query Match 16.5%; Score 335; DB 1; Length 377;
 Best Local Similarity 31.0%; Pred. No. 5.29e-24;

Matches 81; Conservative 54; Mismatches 99; Indels 27; Gaps 22;

Db 33 DACEE-PTPEAMELLIG-KPYEIEGERVDYCKKGYFIPLATHITCDNRHNTLPS 90
 QY 1 EDNELEPPRRNTEILTGNSDQTYPEGTQAIKCRPGYSLGNV-I-MVC-RKGEWALN 57
 Db 91 D-DACYRETCPYIRD-PLNGQAVPANGTYEFGYOMHFIENEGYIIGELITCELKGSVA 148
 QY 58 PLKRCQKRCRPGHGDTPFGFTILGNGVFEYGVKAVYTNEGQYQILGE-INYRE-CDTGG 115
 Db 149 IMSGRPIECYKVLCTPPKIKNGKHTFSEYEV-FEY-L-DAVYSCDPAAGPDPFSLIGE 205
 QY 116 -WTNDIPICEVVKCLPVTAPENCKIYSSAMEPDRYHFGQAVFYCN-S-G---YKIEGD 169

Db 206 STIYCGNSVMSRAAPCEKAVKCRFPVYENGKQISFGKKEFYKATVME--CDGKF-YL 262
 QY 170 EHMCHSDGDFWMSKEKRCVCEISCKSPDYINGSPIS---QKIYKENERFYQKCNMGYEYS 226
 Db 263 DGSDDTYCDNSSTWDPVPVPC 283
 QY 227 ERGDA-VC-TESGM-RPLPSC 244

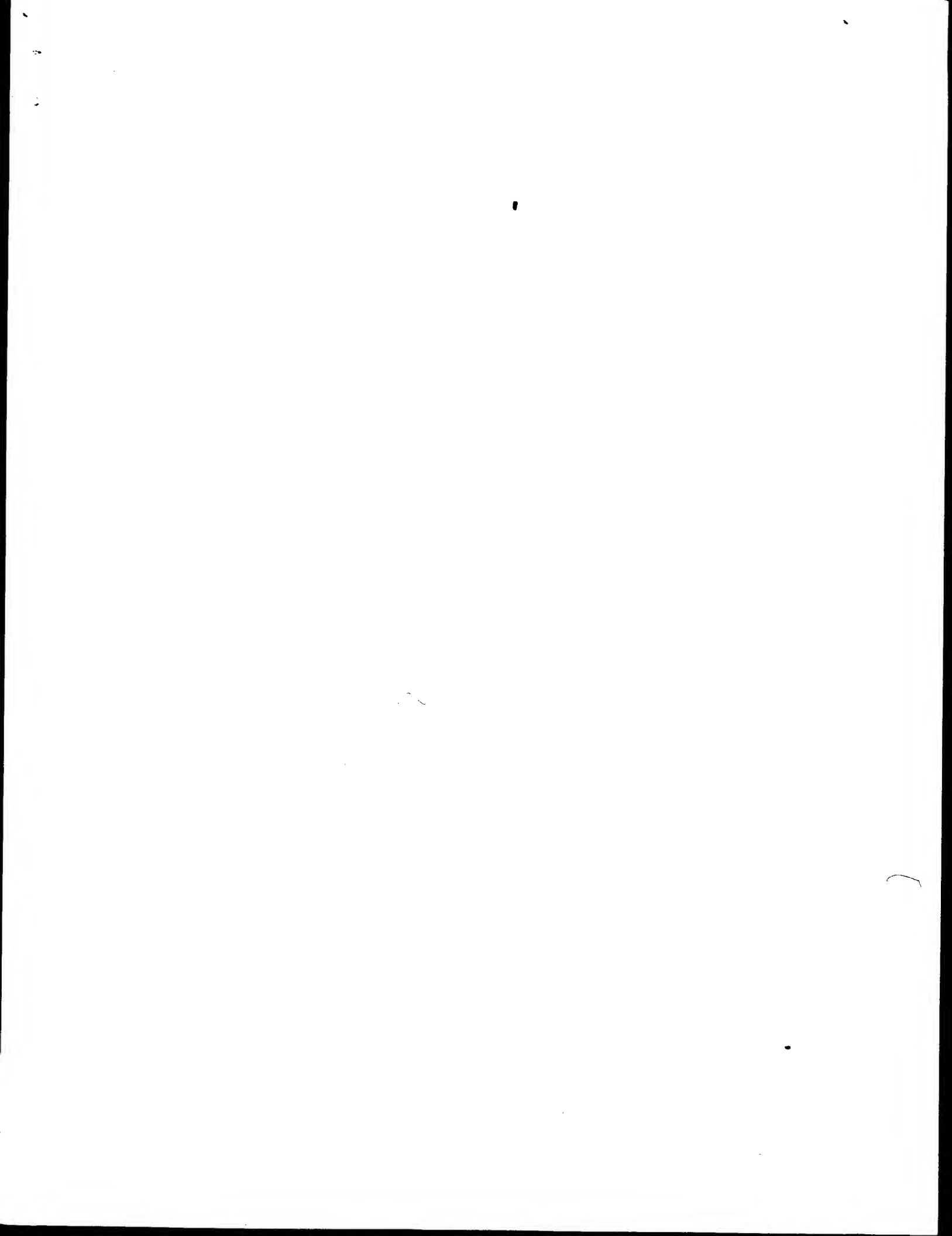
RESULT 13

ID R10924 standard; Protein; 384 AA.
 AC R10924.
 DT 09-MAY-1991 (first entry)
 DE Human membrane cofactor protein isoform.
 KW Mammalian membrane cofactor protein; complement activity;
 KW inflammation; autoimmune diseases; tissue injury.
 FH Key
 FT region
 FT 1..251
 FT /label= short consensus repeats (4)
 FT region
 FT 252-266
 FT /label= serine-threonine rich region, STB
 FT region
 FT 267..280
 FT /label= STC
 FT region
 FT 281..294
 FT /label= unknown region
 FT region
 FT 295..326
 FT /label= hydrophobic region
 FT region
 FT 327..350
 FT /label= CYN2 cytoplasmic tail region
 PN W09102002-A.
 PD 21-FEB-1991.
 PF 20-JUL-1990; U04107.
 PR 21-JUL-1989; US-384210.
 PR 19-APR-1990; US-510709.
 PA (UNITW) UNIV OF WASHINGTON.
 PI Atkinson JP.
 DR WPI: 91-073491/10.
 DR N-PSDB: Q10864.
 PT New recombinant mammalian membrane co-factor protein - for
 PT treating diseases with altered complement activity e.g.
 PT inflammatory and auto-immune conditions
 PS Disclosure: fig 1; 38pp; English.
 CC This human isoform of a membrane cofactor protein (MCP) is useful
 CC in a pharmaceutical compsn. for the treatment of inflammatory and
 CC autoimmune diseases mediated by excess or misdirection of comple-
 CC ment activity, e.g. rheumatoid arthritis or multiple sclerosis.
 CC Protection against tissue injury caused by e.g. myocardial in-
 CC flection or stroke may also be provided. Antibodies raised
 CC against this MCP can be used in autoimmune disease diagnosis to
 CC predict the probability of recurrent miscarriages by testing for
 CC MCP levels in the placenta.
 CC See also Q10865-66 and R10927.
 SQ Sequence 384 AA;

Query Match 16.5%; Score 335; DB 1; Length 384;
 Best Local Similarity 31.0%; Pred. No. 5.29e-24;

Matches 81; Conservative 54; Mismatches 99; Indels 27; Gaps 22;

Db 33 DACEE-PTPEAMELLIG-KPYEIEGERVDYCKKGYFIPLATHITCDNRHNTLPS 90
 QY 1 EDNELEPPRRNTEILTGNSDQTYPEGTQAIKCRPGYSLGNV-I-MVC-RKGEWALN 57
 Db 91 D-DACYRETCPYIRD-PLNGQAVPANGTYEFGYOMHFIENEGYIIGELITCELKGSVA 148
 QY 58 PLKRCQKRCRPGHGDTPFGFTILGNGVFEYGVKAVYTNEGQYQILGE-INYRE-CDTGG 115
 Db 149 IMSGRPIECYKVLCTPPKIKNGKHTFSEYEV-FEY-L-DAVYSCDPAAGPDPFSLIGE 205
 QY 116 -WTNDIPICEVVKCLPVTAPENCKIYSSAMEPDRYHFGQAVFYCN-S-G---YKIEGD 169
 Db 206 STIYCGNSVMSRAAPCEKAVKCRFPVYENGKQISFGKKEFYKATVME--CDGKF-YL 262
 QY 170 EHMCHSDGDFWMSKEKRCVCEISCKSPDYINGSPIS---QKIYKENERFYQKCNMGYEYS 226



Glu-310 was thought to be ignored in translation

```

REFERENCE
#authors Schwaebble, W.; Zwirner, J.; Schulz, T.F.; Linke, R.P.;
#journal Eur. J. Immunol. (1987) 17:1485-1489
#title Human complement factor H: expression of an additional
#cross-references MUID:88053295
#accession A61103
#status not compared with conceptual translation
#molecule_type mRNA
#residues 27-76 #label SC2
#note this is a partial sequence of an alternatively spliced
1.8 kilobase mRNA that is translated to yield a 43 k
form related to factor H

REFERENCE
#authors Sim, R.B.; Discipio, R.G.;
#journal Biochem. J. (1982) 205:285-293
#title Purification and structural studies on the complement-system
control protein beta-1-H (factor H).
#cross-references MUID:83048213
#accession A26505
#molecule_type protein
#residues 19-20,'Q',22-29,'V',31-33,'Q',35 #label SIM

REFERENCE
#authors Barlow, P.N.; Norman, D.G.; Steinkasserer, A.; Horne, T.J.;
#journal Biochemistry (1992) 31:3625-3634
#title Solution structure of the fifth repeat of factor H: A second
example of the complement control protein module.
#cross-references MUID:92232649
#accession S10479
#molecule_type mRNA
#residues 19-20,'Q',22-29,'V',31-33,'Q',35 #label SIM

REFERENCE
#authors Kristensen, T.; Wetzel, R.A.; Tack, B.F.;
#journal J. Immunol. (1986) 136:3407-3411
#title Structural analysis of human complement protein H: homology
with C4b binding protein, beta(2)-glycoprotein I, and the
Ba fragment of B.
#cross-references MUID:86169701
#accession S10479
#molecule_type mRNA
#residues 226-401,'Y',403-449 #label KRI

COMMENT
#cross-references GB:M12383; NID:q180472; PID:AAA52013.1; PID:q180473
Factor H has also been found bound to cell membranes in an unknown
manner. However, it has at least one cell attachment site motif
in repeat 4.
Alternative transcripts of 4.3, 1.8, and 1.4 kilobases are expressed
in liver. See also PIR:NBH0H.

GENETICS
#gene GDB:HP1; HP
#cross-references GDB:1120041; OMIM:134370
#map_position 1q32-1q32
#gene GDB:HP2; HP
#cross-references GDB:129095
#map_position 1q32-1q32

FUNCTION
#description a cofactor in the inactivation of C3b by serine proteinase I;
also increases the rate of dissociation of the C3bBb
complex (C3 convertase) and the (C3b)nb complex (C5
convertase) in the alternative complement pathway
#pathway complement alternate pathway
#supernaturally complement factor H; complement factor H repeat
homology
#alternative splicing; complement alternate pathway;
glycoprotein; plasma

KEYWORDS
#domain signal sequence #status predicted #label SIG\
#product complement factor H, short splice form #status
experimental #label MAR\
#domain complement factor H repeat homology #label FH01\
21-80

```

```

85-141 #domain complement factor H repeat homology #label FH02\
146-205 #domain complement factor H repeat homology #label FH03\
210-262 #domain complement factor H repeat homology #label FH04\
246-248 #region cell attachment (R-G-D) motif\
267-320 #domain complement factor H repeat homology #label FH05\
325-385 #domain complement factor H repeat homology #label FH06\
389-442 #domain complement factor H repeat homology #label FH07\
21-66,'S2-80,85-129,
114-141,146-192,
178-205,210-251,
237-262,267-309,
294-320,325-374,
357-385,389-431,
416-442
#disulfide_bonds #status predicted\
#binding_site carbohydrate (Asn) (covalent) #status
absent

SUMMARY
#length 449 #molecular_weight 51007 #checksum 6077
Query Match 100.0%; Score 2029; DB 1; Length 449;
Best local similarity 100.0%; Pred. No. 0.00e+00;
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 19 EDCNELPPRRNTEILTGSMDSQTYEGTQATYKCRPGYSLGNYIMCRKGWYALNPLR 78
1 EDCNELPPRRNTEILTGSMDSQTYEGTQATYKCRPGYSLGNYIMCRKGWYALNPLR 60
79 KQKRPCCGHPGDPPTGTTLNGVNFEGYKAVYTCNBYGLCEINYRECDTGTGNTDI 138
61 KQKRPCCGHPGDPPTGTTLNGVNFEGYKAVYTCNBYGLCEINYRECDTGTGNTDI 120
Db 139 PICEVVKCLPYTAPENGKIVSAMPDREYHFGQAVFVCSGYKIGDDEMHCSDDGF 198
121 PICEVVKCLPYTAPENGKIVSAMPDREYHFGQAVFVCSGYKIGDDEMHCSDDGF 180
Db 199 SKERPKVEICKSPDYVNGSPISOKIITYKENRFPYKCNMGYSESGDGYVCTESGWR 258
161 SKERPKVEICKSPDYVNGSPISOKIITYKENRFPYKCNMGYSESGDGYVCTESGWR 240
Db 259 LPSCERKSCDNPYIPNGDYSPLRIK 283
241 LPSCERKSCDNPYIPNGDYSPLRIK 265
Qy 241 LPSCERKSCDNPYIPNGDYSPLRIK 265

RESULT 2
ENTRY NBH0H #type complete
TITLE complement factor H precursor, long splice form - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change
22-Jun-1999
ACCESSION S00254; A60238; A54726; A61565; A26505; I72654; S66298
S00254
#authors Ripoche, J.; Day, A.J.; Harris, T.J.R.; Sim, R.B.;
#journal Biochem. J. (1988) 249:593-602
#title The complete amino acid sequence of human complement factor
H.
#cross-references MUID:88134059
#accession S00254
#molecule_type mRNA
#residues 1-1231 #label RIP
#cross-references EMBL:Y00716; NID:q31964; PID:CAA68704.1; PID:q31965
402-Tyr was also found
#note parts of this sequence, including the amino and carboxyl
ends of the mature protein, were confirmed by protein
sequencing

REFERENCE
#authors Estallier, C.; Schwaebble, W.; Dierich, M.; Weiss, E.H.;
#journal Eur. J. Immunol. (1991) 21:799-802
#title Human complement factor H: two factor H proteins are derived
from alternatively spliced transcripts.
#cross-references MUID:91184292
#accession A60238
#status not compared with conceptual translation
#molecule_type mRNA

```



```

959-984,989-1032,
1018-1043,
1048-1091,
1077-1102,
1109-1152,
1138-1163,
1167-1218,
1201-1228
217
529, 802, 822, 882,
911
718,1029,1095
SUMMARY      #length 1231    #molecular-weight 139124    #checksum 9625
Query Match          100.0%; Score 2029; DB 1; Length 1231;
Best Local Similarity 100.0%; Pred. No. 0.0e+00;
Matches   265; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Db      19 EDCNELPPRRNTEITLGSMDQTYEGTQAITYKCRPGRSIGNIYIMCRKGWVAALNPLR 78
|||||
Qy      1 EDCNELPPRRNTIELLGMSDQTYEGTQAITYKCRPRGTRISGANIYIMCRKGWVAALNPLR 60
|||||
Db      79 KCKRPCCGGPCTPRGCTFTLTGCVNFVEYGVAAYVTCNEGYQLGLGEINRYRCDDIDGMTNDI 138
|||||
Qy      61 KCKRPCCGGPCTPRGCTFTLTGCVNFVEYGVAAYVTCNEGYQLGLGEINRYRCDDIDGMTNDI 120
|||||
Db      139 PICEVVKCLPVTAAPNGKIYSSAMEPDREYFEGQAVFYCNCSGKIIEGDDEMCSDDGFW 198
|||||
Qy      121 PICEVVKCLPVTAAPNGKIYSSAMEPDREYFEGQAVFYCNCSGKIIEGDDEMCSDDGFW 180
|||||
Db      199 SKEKRKCYSICKSPDIVINSPISQKITIYENERFOYKCMNGEYSERGDVATESGMWP 258
|||||
Qy      181 SKEKRKCYSICKSPDIVINSPISQKITIYENERFOYKCMNGEYSERGDVATESGMWP 240
|||||
Db      259 LPSCSEKSCDNPIYPINGDISPLRIK 283
|||||
Qy      241 LPSCSEKSCDNPIYPINGDISPLRIK 265
|||||

RESULT      3
ENTRY       NBMSH           #type complete
TITLE       complement factor H precursor - mouse
ALTERNATE_NAMES
ORGANISM    protein beta-1-H
DATE        #format_name Mus musculus #common_name house mouse
            30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change
            22-Jun-1999
ACCESSIONS A26154; 149711; 149728
REFERENCE   A26154
#authors    Kristensen, T.; Tack, B.F.
#journal    Proc. Natl. Acad. Sci. U.S.A. (1986) 83:3963-3967
#title      Murine protein H is comprised of 20 repeating units, 61 amino
            acids in length.
#cross-references MVID:8623353
#accession  A26154
            #molecule_type mRNA
            ##residues     1-1234 ##label KR1
REFERENCE   ##cross-references GB:M12660; NID:g193724; PIDN:AAA37759.1; PID:g387181
#authors    Tatsunuma-Sakai, S.; Nonaka, M.; Nonaka, M.; Harada, Y.
#journal    J. Immunol. (1990) 144:358-362
#title      Demonstration of an unusual allelic variation of mouse factor
            H by the complete cDNA sequence of the H.2 allotype.
#cros-referecnces MVID:90111033
#accession  I49711
            #status       preliminary; translated from GB/EMBL/DBJ
            ##molecule_type mRNA
            ##residues     1-18 ##label RES
            ##cross-references GB:M1979; NID:g193726; PIDN:AAA37762.1; PID:g193729
REFERENCE   I49728
#authors    Munoz-Canoves, P.; Tack, B.F.; Viik, D.P.
```

GENETICS	CLASSIFICATION	KEYWORDS	FEATURE
<p>#journal Biochemistry (1989) 28:9891-9897</p> <p>#title Analysis of complement factor H mRNA expression: Dexanethasone and IFN-gamma increase the level of H in L cells.</p> <p>#cross-references M01D:90148935</p> <p>#accession I49728</p> <p>#status preliminary; translated from GB/EMBL/DBJ</p> <p>#molecule_type mRNA</p> <p>#residues 1-19 #label RE2</p> <p>#cross-references GB:J02891, NID:9193805, PIDN:AAA37795.1, PID:G553926</p> <p>COMMENT Two codominant alleles of factor H are present in mice. Factor H functions as a cofactor in the inactivation of C3b by serine proteinase I and also increases the rate of dissociation of the C3bb complex (C3 convertase) and the (C3b)nb complex (C3 convertase) in the alternative complement pathway.</p>	<p>#map_position 1</p> <p>#superfamily complement factor H; complement factor H repeat</p> <p>homology plasma</p> <p>complement alternate pathway; duplication; glycoprotein;</p>	<p>domain signal sequence #status predicted #label SIG\</p> <p>product complement factor H #status predicted #label MPR\</p>	<p>21-80</p> <p>85-141</p> <p>145-205</p> <p>210-262</p> <p>245-248</p> <p>267-320</p> <p>325-385</p> <p>389-442</p> <p>448-505</p> <p>509-554</p> <p>629-652</p> <p>629-663</p> <p>659-743</p> <p>759-802</p> <p>809-861</p> <p>867-931</p> <p>939-989</p> <p>994-1048</p> <p>1053-1107</p> <p>1114-1168</p> <p>1172-1133</p> <p>21-66,52-80,85-129,114-141,146-192,178-205,210-251,237-262,267-309,294-320,325-374,357-385,389-431,416-442,448-494,477-505,509-553,536-564,569-610,597-622,629-672,658-683,690-732,718-743,752-791,780-802,808-850,836-861,867-920,906-931,936-978,964-989,994-1037,1023-1048,1053-1096,1082-1107,1114-1157,1143-1168,1172-1223,1206-1233,676,721,773,801,1030,1061,1225</p> <p>#disulfide_bonds #status predicted\</p> <p>#binding_site carbohydrate (asn) (covalent) #status predicted</p> <p>#length 1234 #molecular-weight 139081 #checksum 3676</p>

Query Match 70.9%; Score 1439; DB 1; Length 1234;
 Best Local Similarity 67.0%; Pred. No. 0.00e+00;
 Matches 177; Conservative 39; Mismatches 48; Indels 0; Gaps 0;

Db 19 EDCRPPRESEITSGMSSEOLYEGTQATYKCPGRTGTGTYKCKNGKWAASNP8 78
 1 EDCNLPRLRRNTETITGMSDQTYETGTAIKCPGRSLGNVIMCRKGEWALNPLR 60
 Db 79 ICRKRCGHPDTPGSEPLAVGSOFEFGAVYTCDDGTYQLLGEIDYREGADGWINDI 138
 61 KQKRCPCGHPDTPGSEPLAVGSOFEFGAVYTCDDGTYQLLGEIDYREGADGWINDI 120
 Db 139 PLCEVYKCLPTELENGRIVSGAETDQYFVGAVREFGNGEFGKIEGKIEHSCENGLM 198
 121 PLCEVYKCLPTELENGRIVSGAETDQYFVGAVREFGNGEFGKIEGKIEHSCENGLM 180
 Db 199 SNEKRCVETLCTPRVENDGINKVPYKENERHYKCKHGYKPERGDAVCTGSGMS 258
 181 SNEKRCVETLCTPRVENDGINKVPYKENERHYKCKHGYKPERGDAVCTGSGMS 240
 Db 259 QPCEERKRCSPYIINGITTPHRI 282
 241 LPSCERKSCDNPYIPNGDYSPRI 264

RESULT 4
 ENTRY S65551 #type fragment
 TITLE factor H - bovine (fragment)
 ORGANISM #formal_name Bos primigenius taurus #common_name cattle
 DATE 10-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 13-Aug-1999

ACCESSIONS
 REFERENCE S65551
 #authors Soames, C.J.; Day, A.J.; Sim, R.B.
 #journal Biochem. J. (1996) 315:523-531
 #title Prediction from sequence comparisons of residues of factor H involved in the interaction with complement component C3b.
 #cross-references MIMD:96202005
 #accession S65551
 #status preliminary; not compared with conceptual translation
 #molecule_type mRNA
 #residues 1-669 #label SOA
 CLASSIFICATION #superfamily complement factor H; complement factor H repeat homology

FEATURE
 55-114 #domain complement factor H repeat homology #label FHR1
 296-349 #domain complement factor H repeat homology #label FHR2
 355-412 #domain complement factor H repeat homology #label FHR3
 416-471 #domain complement factor H repeat homology #label FHR4
 476-530 #domain complement factor H repeat homology #label FHR5
 538-592 #domain complement factor H repeat homology #label FHR6
 599-651 #domain complement factor H repeat homology #label FHR6
 SUMMARY #length 669 #checksum 8335

Query Match 46.9%; Score 952; DB 2; Length 669;
 Best Local Similarity 64.3%; Pred. No. 9.39e-211;
 Matches 117; Conservative 29; Mismatches 36; Indels 0; Gaps 0;

Db 7 LAENQFYGAKVYTCDEGQYMGVGNFRCDTNGTNDIPCEVYKCLPYTEPENGKI 66
 80 LTGQNVFYGKAVYTCDEGQYMGVGNFRCDTNGTNDIPCEVYKCLPYTEPENGKI 139
 Db 67 FSDALEPQOETYGQVQVQECNSGYMLDGPQKIHCSAGVSAETPCVETLFCAPVILN 126
 140 VSSAMEPQOETYGQVQVQECNSGYMLDGPQKIHCSAGVSAETPCVETLFCAPVILN 199
 Db 127 GOAVLPKATYQNERVQRCAGFEYGGGRTVCTKSGMTAPRCIELTCDPPRIPIGVY 186
 200 GSPISQKIITYKENERFOYKCMNGEYSEKGDVACTESGMRPLPSCERKSCDNPYIPNGDY 259
 Db 187 RP 188

QY 260 SP 261

RESULT 5
 ENTRY S46199 #type complete
 TITLE probable complement regulatory plasma protein SBI - barred sand bass
 ORGANISM #formal_name Paralichthys nebulifer #common_name barred sand bass
 DATE 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 13-Aug-1999

ACCESSIONS
 REFERENCE S46199
 #authors Dahmen, A.; Kaidoh, T.; Ziefel, P.F.; Gigli, I.
 #journal Biochem. J. (1994) 301:391-397
 #title Cloning and characterization of a cDNA representing a putative complement-regulatory plasma protein from barred sand bass (Paralichthys nebulifer).
 #cross-references MIMD:94318039
 #accession S46199
 #molecule_type mRNA
 #residues 1-1053 #label DAH1
 #experimental_source liver

GENETICS
 SBI
 CLASSIFICATION #superfamily complement factor H repeat homology
 KEYWORDS glycoprotein
 FEATURE 89-145
 334-389 #domain complement factor H repeat homology #label FH01
 450-502 #domain complement factor H repeat homology #label FH02
 569-624 #domain complement factor H repeat homology #label FH03
 682-738 #domain complement factor H repeat homology #label FH04
 743-802 #domain complement factor H repeat homology #label FH05
 935-989 #domain complement factor H repeat homology #label FH06
 993-1052 #domain complement factor H repeat homology #label FH06
 SUMMARY #length 1053 #molecular_weight 11737 #checksum 8482

Query Match 28.9%; Score 587; DB 2; Length 1053;
 Best Local Similarity 37.1%; Pred. No. 3.83e-116;
 Matches 91; Conservative 41; Mismatches 101; Indels 12; Gaps 9;

Db 47 EASYDGRQVRYGCVNGY-S-GFEKVCESGKWTGCA--KQDRSCGHPDQAFADFL 102
 21 DQTEGEGQAIYKCRPGYSLGNVIMCRKGEWALNPLRCKQKPCGHPDTPFGFTL 80
 Db 103 AGNDPFGSKVYTCQKQYMGVGNFRCDTNGTNDIPCEVYKCLPYTEPENGKI 161
 81 TCGNVFYGKAVYTCDEGQYMGVGNFRCDTNGTNDIPCEVYKCLPYTEPENGKI 140
 Db 162 ICG--PE-EATRGVNVFRSCKSRSEILDSPELYCDERGDMSGVPKCAITCAIPIEN 218
 141 SSAMEPQOETYGQVQVQECNSGYMLDGPQKIHCSAGVSAETPCVETLFCAPVILN 199
 QY 219 GNVGAIIRYKENDVLRHCDRAFKHIDPSTCIQKIAKMSPPPLCESIKCRTINDG 278
 200 GSPISQKIITYKENERFOYKCMNGEYSEKGDVACTESGMRPLPSCERKSCDNPYIPN 256
 Db 279 TRPE 283
 QY 257 GDYSP 261

RESULT 6
 ENTRY S53711 #type complete
 TITLE CABP alpha chain precursor - rabbit
 ORGANISM #formal_name Oryctolagus cuniculus #common_name domestic rabbit
 DATE 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999


```

ACCESSIONS      S53711
REFERENCE        de Frutos, P.G.; Dahlbaeck, B.
#authors        Blochim, Biophys. Acta (1995) 1261:285-289
#journal        CDNA structure of rabbit C4b-binding protein alpha-chain.
#title          Preserved sequence motif in complement regulatory protein
                  modules which bind C4b.
#cross-references  PubMed:95226458
#accession      S53711
#status        Preliminary; nucleic acid sequence not shown
#molecule_type mRNA
#residues      1-597 ##label DEF
#cross-references  EMBL:Z35490
CLASSIFICATION  #superfamily C4b-binding protein alpha chain: complement
                  factor H repeat homology
FEATURE
50-107          #domain complement factor H repeat homology #label FH1
112-169         #domain complement factor H repeat homology #label FH2
174-234         #domain complement factor H repeat homology #label FH3
239-294         #domain complement factor H repeat homology #label FH4
299-360         #domain complement factor H repeat homology #label FH5
364-422         #domain complement factor H repeat homology #label FH6
426-480         #domain complement factor H repeat homology #label FH7
484-538         #domain complement factor H repeat homology #label FH8
SUMMARY         #length 597 #molecular-weight 66130 #checksum 6473

Query Match      20.0%; Score 406; DB 1; Length 597;
Best Local Similarity 32.9%; Pred. No. 1,11e-70;
Matches 83; Conservative 49; Mismatches 97; Indels 23; Gaps 19;

Db
66 SENEYOTGTLKTKYRFRNGINPLTCKRGLW--SYDTF--CYAKRRNEDDLNQG 122
| : | | | | | | | | | | | | | | | | | | | | | | | | | | |
20 SDQTYPGTAIKKCRGYSLSG-NVIMVCR-KGEWVALLPKRCQKRCRPGHDPDPFGT 77
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
123 VEKRT-D-SEFSQGLESCSEGYLLIGSTT-SHCDIOEKGVESDPLKCEIYVCEPPN 179
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
78 FTLTGNVEFYGVKAYVTCNEGYYQLLGEINRYRCDT-D-G--WTNDIPICEVKKCLPYTA 133
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
180 IINGK-HNGENE-DIH-TYGSVTVYSCNPRFSLGEMISICTYKNTVGVSSSPYCKE 236
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
134 PENGKIVSSAMPEDRYHFGQAVRFVCSNGYKIEGDEMHCS--DD--GPMSEKPKCIVE 189
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
237 IICSPNPNPHKGIISGFGPIYKDSIMYTCIDGFYLRGSLIHCELDSEWNPSPVCS 296
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
190 ISCKSPDIVNGSPISQ-KIITYKENERFYCKNMGYESENDAVCI-ESGNRLPDS-CEE 246
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db
297 NSCLGLPNVPHA 308
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 247 KSC-DNPIYIPNG 257
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 7
ENTRY      MNVZSP
TITLE      apolipoprotein H homolog precursor - vaccinia virus
ALTERNATE_NAMES 3K secretory protein; C3J protein; virokin
ORGANISM   #formal_name vaccinia virus
#note      host Homo sapiens (man)
          31-Dec-1989 #sequence_revision 30-Jun-1990 #text_change
          22-Jun-1999
ACCESSIONS A31005; B42504
REFERENCE
#authors   Kotwal, G.U.; Moss, B.
#journal   Nature (1988) 335:176-178
#title     Vaccinia virus encodes a secretory polypeptide structurally
          related to complement control proteins.
#cross-references  PubMed:88318974
#accession      A31005
#molecule_type DNA
#residues      1-263 ##label KOT
#cross-references  GB:X13166; NID:g60690; PIDN:CAA31564.1; PID:g60691
#experimental_source strain WR
REFERENCE      A42501
#authors      Goebel, S.J.; Johnson, G.P.; Perkus, M.E.; Davis, S.W.;

```

```

#journal        Winslow, J.P.; Paolletti, E.
#title          Virology (1990) 179:517-563
#accession      Appendix to "The complete DNA sequence of vaccinia virus".
#molecule_type DNA
#residues      1-263 ##label GOE
#cross-references  GB:M35027; NID:g335317; PIDN:AAA47997.1; PID:g335345
#experimental_source strain Copenhagen
REFERENCE      A42531
#authors      Goebel, S.J.; Johnson, G.P.; Perkus, M.E.; Davis, S.W.;
#journal        Winslow, J.P.; Paolletti, E.
#title          Virology (1990) 179:247-266
#cross-references  The complete DNA sequence of vaccinia virus.
#note          annotation: possible protein-coding frames
          neither amino acid nor nucleotide sequence is given
          #superfamily herpesvirus complement control protein;
          complement factor H repeat homology
          duplication; extracellular protein
CLASSIFICATION
KEYWORDS
FEATURE
1-19         #domain signal sequence #status predicted #label SIG
20-263       #product C4b-binding protein homology #status predicted
          #label MAT
21-81        #domain complement factor H repeat homology #label FH1
86-143       #domain complement factor H repeat homology #label FH2
148-201      #domain complement factor H repeat homology #label FH3
206-261      #domain complement factor H repeat homology #label FH4
SUMMARY      #length 263 #molecular-weight 28629 #checksum 8152

Query Match      19.1%; Score 388; DB 1; Length 263;
Best Local Similarity 34.2%; Pred. No. 2.94e-66;
Matches 80; Conservative 37; Mismatches 96; Indels 21; Gaps 17;

Db
40 ANANYNGITIEYLCPYRKQKMPYIAKCTGTG-TT--FNQCIKRCPSPRDINQ 96
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
20 SDQTYPGTAIKKCRGYS--LGNVIMVCRKGEWVALLPKRCQKRCRPGHDPDPFGT 77
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
97 LDI-GG-V-DFGSSITYSNGSYHLIGESKSYCELGSTGSMWNPAIPCESVKCSPPS 153
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
78 FTLTGNVEFYGVKAYVTCNEGYYQLLGEIN-YRE-CDTGG--WTNDIPICEVKKCLPYTA 133
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
154 ISNGR--HNGYE-DF-YTDSVTVYSCNNGSYSLIGNSGYLCSGGE-WS-DPTCOIVKCP 207
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
134 PENGKIVSSAMPEDRYHFGQAVRFVCSNGYKIEGDEMHCSDDGPMSEKPKCIVEISCK 193
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db
208 HPTISNGYLSGFKRSYSYNDVDFKCKYGYKLSGSSSTCPCGNTWPKELPKC 261
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 194 SPDVINGSPISQ-KIITYKENERFYCKNMGYESENDAVCTESG-WRP-LPSC 244
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 8
ENTRY      NBHUC4
TITLE      C4b-binding protein alpha chain precursor - human
ALTERNATE_NAMES C4BP; proline-rich protein
ORGANISM   #formal_name Homo sapiens #common_name man
#note      13-Aug-1986 #sequence_revision 30-Jun-1993 #text_change
          22-Jun-1999
ACCESSIONS A33568; S02372; A90326; A24182; A93134; S29492; A31785;
          I52244; A03210
REFERENCE
#authors      Matsuguchi, T.; Okamura, S.; Aso, T.; Sata, T.; Niho, Y.
#journal      Biochem. Biophys. Res. Commun. (1989) 165:138-144
#title        Molecular cloning of the cDNA coding for proline-rich protein
          (PRP): identity of PRP as C4b-binding protein.
#cross-references  PubMed:90073699
#accession      A33568
#molecule_type mRNA
#residues      1-597 ##label MA1
#cross-references  GB:M31452; NID:g190501; PIDN:AAA36507.1; PID:g190502
#note          the authors translated the codon GGA for residue 492 as
          Glu
REFERENCE      S02372
#authors      Lintin, S.J.; Lewin, A.R.; Reid, K.B.M.

```

	REFERENCE	#note	This peptide appears to bind protein S
	#residues	495-505,'X','Y','507'-510,'X','512-515	#label Suz
	#authors	Ahlback, B.; Smith, C.A.; Muller-Eberhard, H.J.	
	#journal	Proc. Natl. Acad. Sci. U.S.A. (1983)	80:3461-3465
	#title	Visualization of human C4b-binding protein and its complexes with vitamin K-dependent protein S and complement protein C4b.	
	#cross-references	MUID:83221615	
	#contents	annotation; electron microscopy; three-dimensional structure; ligand binding	
	REFERENCE	I52244	
	#authors	Aso, T.; Okamura, S.; Matsuguchi, T.; Sakamoto, N.; Sata, T.; Niho, Y.	
	#journal	Biochem. Biophys. Res. Commun. (1991)	174:222-227
	#title	Genomic organization of the alpha chain of the human C4b-binding protein gene.	
	#cross-references	MUID:91113199	
	#accession	I52244	
	#status	translation not shown; translated from GB/EMBL/DDBJ	
	#molecule-type	DNA	
	#residues	1-597	#label ASO
	#cross-references	GB:M62486; NID:g190498; PIDN:AAA36506.1; PID:g190500	
	COMMENT	C4b controls the classical pathway of complement activation. It binds as a cofactor to C3b/C4b inactivator (C3BINA), which then hydrolyzes the complement fragment C4b. It also accelerates the degradation of the C4bC2a complex (C3 convertase) by dissociating the complement fragment C2a.	
	COMMENT	C4BP occurs in plasma in two forms, both of which bind complement fragment C4b. The prevalent higher molecular weight form contains 7 alpha chains and one beta chain, which are linked by disulfide bonds. The beta chain binds the vitamin K-dependent plasma protein S. A minor form lacks the beta chain. Bound protein S is inactive as a cofactor for protein C inactivation of coagulation factors V and VIII.	
	COMMENT	The molecule has a central body supporting seven tentacles (alpha chains), each with the binding site for C4b at the peripheral end.	
	GENETICS	GDB:C4BPA	
	#gene	#cross-references GDB:120568; OMIM:120830	
	#map_position	Iq32-Iq32	
	#introns	48/1; 110/1; 143/2; 172/1; 236/1; 297/1; 362/1; 425/1; 482/1; 540/3	
	COMPLEX CLASSIFICATION	octamer of seven alpha chains and one beta chain #superfamily C4b-binding protein alpha chain; complement factor H repeat homology acute phase; chylomicron; complement pathway; duplication; glycoprotein; plasma	
	KEYWORDS		
	FEATURE	#domain signal sequence #status predicted #label SIG\	
	1-48	#product C4b-binding protein alpha chain #status predicted #label MAT\	
	49-597		
	50-108	#domain complement factor H repeat homology #label FH1\	
	113-170	#domain complement factor H repeat homology #label FH2\	
	175-334	#domain complement factor H repeat homology #label FH3\	
	239-294	#domain complement factor H repeat homology #label FH4\	
	299-360	#domain complement factor H repeat homology #label FH5\	
	364-422	#domain complement factor H repeat homology #label FH6\	
	381-404	#region complement C4b binding #status predicted\	
	426-480	#domain complement factor H repeat homology #label FH7\	
	446-538	#domain complement factor H repeat homology #label FH8\	
	221-506, 528	#binding_site carboxyrate (Asn) (covalent) #status experimental	
	SUMMARY	#length 597 #molecular-weight 67033 #checksum 6374	
	Query Match	18.8%; Score 382; DB 1; Length 597;	
	Best Local Similarity	31.5%; Pred. No. 8,65e-65;	
	Matches	79; Conservative 52; Mismatches 98; Indels 22; Gaps 20;	
Db	67	TETRTKGTTLTKYLCPDPYVRSHSTQTLLNSDEWV-VNTFCIT-KR-CRAPGELRNQO	123
:	:	: : : : : : : : :	:
QY	20	SPQTPPEPGTALIKCRPGY-RSLGNVIWCKR-GEWALNPILKKCRKPCGHGADPFGT	77

Db 124 VEIKT-DL-SFGSQIEFSCSEGFLLIGSTSRCEVDGKMSHPDCEIYKCKPPDI 181
 QY 78 FTLLGNGVFEYGVAAVWCMNEGIOILLGINTNR-EC-DTD-GWTDIPICEVYKCLPVTAP 134
 Db 182 RNGR-HSG-E-ENFYAVGFSVTYSCDPRFSLILGHASISCTVENETIGVMPSPPTCEKI 237
 QY 135 ENGRIVSSAMEPDRHYHGQAVRFVNCNGYKIEGDEMHCS-D-D-GFWSKEKPKCEI 190
 Db 238 TCRKPDVSHGMSVGFQIYKQDIYKCKCKGVLKRSSTYIHCDDAKNPSPPACEPN 297
 QY 191 SCRPDVIYNSPISQ-KITIKENERFOYKCMWGYEYSEKDAVC-TESGMRP-LPSCEEK 247
 Db 298 SCINLPDIPHA 308
 QY 248 SCDN-PYIPNG 257

ENTRY 9
 TITLE NEMS
 ALTERNATE_NAMES #type complete
 50K serum glycoprotein; activated protein C-binding protein;
 beta-2-glycoprotein I
 ORGANISM #formal_name Mus musculus #common_name house mouse
 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change
 28-May-1999
 ACCESSIONS A43286; J02243
 REFERENCE A43286
 #authors Nonaka, M.; Matsuda, Y.; Shirosaki, T.; Moriwaki, K.; Nonaka, M.; Natsunne-Sakai, S.
 #journal Genomics (1992) 13:1082-1087
 #title Molecular cloning of mouse beta-2-glycoprotein I and mapping of the gene to chromosome 11.
 #cross-references MVID:92372000
 #accession A43286
 #molecule_type mRNA
 #residues 1-253, 'A', 254-277, 'N', 279-345 #label NON
 #cross-references GB:D10056
 #note The authors translated the codon ACT for residue 253 as Ala and ATG for residue 278 as Asn
 #note sequence extracted from NCBI backbone (NCBIN:111791, NCBI:P:111794) and corrected to correspond with the nucleotide translation

REFERENCE J02243
 #authors Sellar, G.C.; Steel, D.M.; Zafiroopoulos, A.; Seery, L.T.; Whitehead, A.S.
 #journal Biochem. Biophys. Res. Commun. (1994) 200:1521-1528
 #title Characterization, expression and evolution of mouse beta2-glycoprotein I (apolipoprotein H).
 #cross-references MVID:94242017
 #accession J02243
 #molecule_type mRNA
 #residues 1-251, 'R', 253-345 #label SEL
 #cross-references GB:S70439; NID:g546780; PID:AA30789.1; PID:g546781
 #experimental_source liver

GENETICS B2gpi
 #gene map_position 11
 CLASSIFICATION #superfamily apolipoprotein H; complement factor H repeat
 #homology
 #chylomicron; duplication; glycoprotein; HDL; heparin binding; lipid binding; monomer; plasma; VLDL

KEYWORDS
 #domain signal sequence #status predicted #label SIG
 #product apolipoprotein H #status predicted #label MAT
 #domain complement factor H repeat homology #label FH1
 #domain complement factor H repeat homology #label FH2
 #domain complement factor H repeat homology #label FH3
 #domain complement factor H repeat homology #label FH4
 #domain complement factor H repeat homology #label FH5
 #domain complement factor H repeat homology #label FH6
 #domain complement factor H repeat homology #label FH7
 #domain complement factor H repeat homology #label FH8

FEATURE
 1-19
 20-345
 23-76
 84-137
 142-200
 205-260
 264-325
 23-66, 51-79, 84-124,
 110-137, 142-188,
 174-200, 205-248,

234-260, 264-315,
 300-307, 325-345
 105, 117, 162, 183,
 193
 #disulfide_bonds #status predicted\
 #binding site carbohydrate (Asn) (covalent) #status predicted

SUMMARY
 #length 345 #molecular_weight 38618 #checksum 2016

Query Match 18.3%; Score 371; DB 1; Length 345;
 Best Local Similarity 29.4%; Pred. No. 4, 20e-62;
 Matches 74; Conservative 49; Mismatches 110; Indels 19; Gaps 15;

Db 40 SYDPEQIVYSCRPYRSGMRFRFCPLTGMV-PINTLR-CYPRVCPPFGLLENGIVRY 97
 QY 23 TYPETGTAIKCRPGRISLGNVTK-VCR-KGEWALNPLKRCCKRPGHGDPPFGFTL 80
 Db 98 TS--FEYPRNISFACNPGF-FLNGTSSSKCTEGKMSPDIPACARITCPPVPKFA 153
 QY 81 TGGVFEYGVKAVYTCNEGQYLLGEIYRCDTDG-WTNDIPICEVYKCLPYAPENKGI 139
 Db 154 KDIRPSAGNSLXODIYVFCLEPHFAMIGNDIYMCIEQGMWPL-PECLEVKCPFPPRE 212
 QY 140 VSSAMEPDRHYHGQAVRFVNCNGYKIEGDEMHCSDDGFWSKKPKCEIISCK-SPDVI 198
 Db 213 NGYVNPAPKPVLLIKDKATPG--CHETFKLDGPEAECTRTGWSPLPCRE-SCKLPRV 269
 QY 199 NG--S-PISKIITKENENEROYKCMWGYEYSEKDAVCTESG-WRPLPSCEEKSCDPYI 254
 Db 270 KATVLYIQMRVK 281
 QY 255 PNGD-YSPULRIK 265

RESULT 10
 ENTRY S57953
 TITLE C4b protein alpha chain precursor - rat
 ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
 DATE 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change
 20-Aug-1999

ACCESSIONS S57953
 REFERENCE S57953
 #authors Hillarp, A.; Thern, A.; Dahlback, B.
 #submission Submitted to the EMBL Data Library, July 1995
 #description Molecular cloning of rat C4b-binding protein alpha- and beta-chains: structural and functional relationships among human, bovine, rabbit, mouse and rat proteins.

#accession S57953
 #status preliminary
 #molecule_type mRNA
 #residues 1-558 #label HIL
 #cross-references EMBL:250051; NID:g899379; PID:CAA90391.1; PID:g899380

CLASSIFICATION #superfamily C4b-binding protein alpha chain; complement factor H repeat homology

FEATURE
 15-72
 77-134
 139-199
 204-258
 263-324
 328-386
 390-443
 447-501
 #domain complement factor H repeat homology #label FH1
 #domain complement factor H repeat homology #label FH2
 #domain complement factor H repeat homology #label FH3
 #domain complement factor H repeat homology #label FH4
 #domain complement factor H repeat homology #label FH5
 #domain complement factor H repeat homology #label FH6
 #domain complement factor H repeat homology #label FH7
 #domain complement factor H repeat homology #label FH8

SUMMARY
 #length 558 #molecular_weight 62265 #checksum 9928

Query Match 18.3%; Score 371; DB 2; Length 558;
 Best Local Similarity 30.5%; Pred. No. 4, 20e-62;
 Matches 73; Conservative 57; Mismatches 87; Indels 22; Gaps 18;

Db 43 YNCRPGYRASSQSILXCKPLGKW-QIN-IA-CVKKSCRNPGLDONGKVEYKT-D-FLFG 97
 QY 32 YKCRPGY-RSLGNVIMCRK-GEWVALNPLKRCCKRPGHGDIPFGFTLLTGGNVFEYG 89
 Db 98 SQIEFSCSEGITLLIGSISTYCEIOGKGVSWSDPLPECVIACGMPPDIS-NGK--HNGRE 154

```

OY 90 KVAATTCBEGTOLDE-INTRECDTGD--WTNDIPICEVVK-LPTAPENKIVSSAME 145
Db 155 EEF-FTYRSTVYKCDPFTLLGNASITCTVANKTVGVWSPSPPCERICRPMVLTGT 213
OY 146 PDREYHFGQAVRFVNSGKIKGDEMHCS--D--GFMSKRPKPCVEISCSPDINIS 201
Db 214 INSGKHKYKKKSDRFPVCGKGFVLRGSGVHCEADGSKMSPVVCELSCTDIPDIPNA 272
OY 202 PISO-KIITYKENRFOYKCNMGYEYSEKGDVAVC-TESGWRPLPSCSEKSC-DNPTIPNG 257

RESULT 11
ENTRY 11
TITLE C36838 #type complete
ALTERNATE_NAMES D12L protein
ORGANISM #formal_name variola virus
DATE 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
ACCESSIONS C36838
REFERENCE A36859
#authors Blinov, V.M.
#submission submitted to Genbank, November 1992
#description not shown.
#accession C36838
#molecule_type DNA
#residues 1-263 #label BLI
#cross-references GB:X69198; NID:9456758; PIDN:CAA48953.1; PID:9297195
CLASSIFICATION #superfamily herpesvirus complement control protein;
complement factor H repeat homology

FEATURE
21-81
86-143 #domain complement factor H repeat homology #label FH1\
148-201 #domain complement factor H repeat homology #label FH2\
206-261 #domain complement factor H repeat homology #label FH3\
SUMMARY #length 263 #molecular_weight 28789 #checksum 8771

Query Match 17.6%; Score 357; DB 1; Length 263;
Best Local Similarity 33.6%; Pred. NO. 1,06e-58;
Matches 79; Conservative 38; Mismatches 95; Indels 23; Gaps 19;

Db 40 ANANNIGDTIEVLCLEPKRKMGPIYACTGTG-TL-FNOCIRKRCPSRDINGH 96
OY 20 SDQTYEGGTQAIKCRGRYS--LGNVIMCRKGEMALNPRLRCQKRPCHPDTPFGT 77
Db 97 LDI-IGG-V-DEGSSITYSCSGYLLIEYKSYCKLGSTGSMWNPAPICESVQDPPS 153
OY 78 FTLTGGVFEYGVKAVYTCNEGQYQLGEIN-Y-RECDTGD--WTNDIPICEVVK-LPVT 132
Db 154 IS-NGR--HNGYN-DF-YTDSVVTYSCNSGYSLIGNSGYLCSGE-WSPN-PTQIYKC 206
OY 133 APENGRIVSAMPDREYHFGQAVRFVNSGKIKGDEMHCSDDGWSKRPKVEISC 192
Db 207 PHFTIINGYISSGFKSYSDNVVFTCKYGYLLSGSSSTGSPGNTWQDELPRC 261
OY 193 KSPDIVINGSPISO-KIITYKENRFOYKCNMGYEYSEKGDVAVC-WRP-LPSC 244

RESULT 12
ENTRY 12
TITLE T16833 #type complete
ORGANISM #formal_name Caenorhabditis elegans
DATE 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
ACCESSIONS T16833
REFERENCE T18586
#authors Gettsel, C.
#submission submitted to the EMBL Data Library, April 1996
#description The sequence of C. elegans cosmid T07H6.
#accession T16833
#status Preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA

```

```

#residues 1-560 #label GEI
#cross-references EMBL:U53344; NID:91255886; PID:91255889;
#experimental_source strain Bristol N2; clone T07H6
GENETICS
#gene CESP:T07H6.5
#map_position 10
#introns 14/1: 75/1: 102/3: 128/1: 186/1: 272/2: 326/1: 422/1:
475/1: 527/1
SUMMARY #length 560 #molecular_weight 61619 #checksum 2418

Query Match 17.4%; Score 353; DB 2; Length 560;
Best Local Similarity 32.2%; Pred. No. 9.92e-58;
Matches 68; Conservative 42; Mismatches 82; Indels 19; Gaps 15;

Db 117 AQWGPD-LR-CKAPACDPDGIENG--LREGDTFEYHHYKSCNPGFLVGSST-RQ 170
OY 51 GEWALNPLRRCQKRPCHPDTPFGTLLGNVFEYGVKAVYTCNEGQYQLGEINRE 110
Db 171 CSNGETNEPANCATKATCSRPSSPLHGKVVGSSL-T---YQ-S-VYIYSCDHGRVYG 224
OY 111 CDIDG-WTNDIPICEVVKCLVYTAPEKNTIVSAMPDREYHFGQAVRFVNSGKIKG 169
Db 225 VQRLCLAGIINGNEPRCEETRCGYLPTLPNGYTESSETSGAVALFRCLETMTHE-GA- 282
OY 170 EEMHCSDGGEFWSKRPKVEISCKS-PDIVINGSPISO-KIITYKENRFOYKCNMGYEYSE 228
Db 283 SKAKMEDGQWAPILPRLA-SCRYPHONG 312
OY 229 GDAVCTESG-WR-PLPSCSEKSCDNPTIPNG 257

RESULT 13
ENTRY 13
TITLE A30359 #type complete
ALTERNATE_NAMES P-selectin precursor - human
ORGANISM CD62 antigen; granule membrane protein 140
DATE 18-Oct-1989 #sequence_revision 30-Sep-1991 #text_change 29-Aug-1997
ACCESSIONS A30359
REFERENCE A30359
#authors Johnston, G.I.; Cook, R.G.; McEver, R.P.
#journal Cell (1989) 56:1033-1044
#title Cloning of GMP-140, a granule membrane protein of platelets and endothelium: sequence similarity to proteins involved in cell adhesion and inflammation.
#cross-references MIMD:89168432
#accession A30359
#molecule_type mRNA
#residues 1-830 #label JOH
#cross-references GB:M25322
#note parts of this sequence, including the amino end of the mature protein, were confirmed by protein sequencing

GENETICS
#gene GDB:SELP: GRNP
#map_position 1422-1925
CLASSIFICATION #superfamily complement factor H repeat homology; EGF
homology
KEYWORDS cell adhesion; glycoprotein; phosphohistidine;
phosphoprotein; surface antigen; transmembrane protein

FEATURE
1-41
42-830 #domain signal sequence #status predicted #label SIG\
163-194 #product P-selectin #status experimental #label MAR\
200-257 #domain EGF homology #label EGF\
324-381 #domain complement factor H repeat homology #label FH01\
386-443 #domain complement factor H repeat homology #label FH02\
448-505 #domain complement factor H repeat homology #label FH03\
510-567 #domain complement factor H repeat homology #label FH04\
572-629 #domain complement factor H repeat homology #label FH05\
642-699 #domain complement factor H repeat homology #label FH07\

```


Fri Jun 9 10:53:38 2000

US-09-316-163-10.rpt

Page 11

QY 196 DVLNG-SPISOKIYYKENERFOYKCNMGE-YSERGDVCTESGWRP-LPSC 244

Search completed: Thu Jun 8 21:41:41 2000
Job time : 20 secs.

(T.M.)

Distribution rights by Oxford Molecular Ltd

ot generated

1 EDCNE

PAM 150
Car 11

83857 seqs, 30454973 residues

Listing first 45 summaries

1:swtssprot

Mean 43.808; Variance 61.409; scale 0.713

ived by analysis of the total score distribution.

SUMMARIES

Query	Match	Length	DB	ID	Description	Pred. No
00.0	1231	1	CEAH_HUMAN	COMPLEMENT FACTOR H PR	0.00e+00	
70.9	1234	1	CFAH_MOUSE	COMPLEMENT FACTOR H PR	0.00e+00	
19.1	263	1	VCP_VACCV	COMPLEMENT CONTROL PRO	7.02e-77	
18.8	558	1	C4BP_HUMAN	C4B-BINDING PROTEIN AL	3.17e-77	
18.3	345	1	APOH_MOUSE	BETA-2-GLYCOPROTEIN I	3.37e-77	
18.3	558	1	C4BP_RAT	C4B-BINDING PROTEIN AL	3.37e-77	
17.0	830	1	LEM3_HUMAN	P-SELECTIN PRECURSOR (8.01e-6	
16.8	360	1	CCPB_HGSA	COMPLEMENT CONTROL PRO	5.23e-6	
16.8	468	1	C4BP_MOUSE	C4B-BINDING PROTEIN PR	5.23e-6	
16.8	768	1	LEM3_RAT	P-SELECTIN PRECURSOR (5.23e-6	
16.7	610	1	C4BP_BOVIN	C4B-BINDING PROTEIN AL	3.41e-6	
16.7	768	1	LEM3_MOUSE	P-SELECTIN PRECURSOR (1.83e-6	
16.5	345	1	APOH_BOVIN	BETA-2-GLYCOPROTEIN I	4.13e-6	
16.5	377	1	MCP_HUMAN	MEMBRANE COFACTOR PROT	2.22e-6	
16.4	345	1	APOH_HUMAN	BETA-2-GLYCOPROTEIN I	7.71e-6	
16.4	2039	1	CRL_HUMAN	COMPLEMENT RECEPTOR	1.73e-5	
15.6	610	1	LEM2_HUMAN	E-SELECTIN PRECURSOR (1.59e-5	
15.5	612	1	LEM2_MOUSE	E-SELECTIN PRECURSOR (1.01e-5	
15.3	958	1	HIG_DROME	LOCOMOTION-RELATED PRO	1.19e-5	
15.2	484	1	LEM2_PIG	E-SELECTIN PRECURSOR (4.09e-5	
15.1	345	1	APOH_CANFA	BETA-2-GLYCOPROTEIN I	1.40e-5	
15.1	551	1	LEM2_RABIT	E-SELECTIN PRECURSOR (1.40e-5	
14.7	549	1	LEM2_RAT	E-SELECTIN PRECURSOR (1.03e-5	

24	293	4.4	381	1	DAF_HUMAN	COMPLEMENT	DECAY-ACCEL	4.0e-49
25	288	14.2	769	1	DAF_SHEEP	P-SELECTIN	PRECURSOR (8.55e-48
26	284	14.0	340	1	DAF_FONPY	COMPLEMENT	DECAY-ACCEL	9.72e-47
27	283	13.9	611	1	DAF_CAFPA	E-SELECTIN	PRECURSOR (1.78e-46
28	280	13.8	507	1	DAF_CAVDO	COMPLEMENT	DECAY-ACCEL	1.10e-45
29	276	13.6	297	1	APHO_RAT	BETA-2-GLYCOPROTEIN I	1.24e-44	1.24e-44
30	276	13.6	485	1	LEM2_BOVIN	L-E-SELECTIN	PRECURSOR (1.44e-44
31	276	13.6	1019	1	LFC_TACR	LIMBUS CLOTTING FACTO	1.24e-44	1.24e-44
32	270	13.3	1033	1	CR2_HUMAN	COMPLEMENT	RECEPTOR TY	4.62e-44
33	262	12.9	668	1	F13B_MOUSE	COAGULATION FACTOR XII	5.63e-40	6.15e-40
34	258	12.7	661	1	F13B_HUMAN	COAGULATION FACTOR XII	6.15e-40	6.15e-40
35	257	12.7	1025	1	CR2_MOUSE	COMPLEMENT	RECEPTOR TY	1.12e-39
36	254	12.5	390	1	DAF1_MOUSE	COMPLEMENT	DECAY-ACCEL	6.67e-39
37	254	12.5	646	1	LEM3_BOVIN	P-SELECTIN	PRECURSOR (6.67e-39
38	238	11.7	407	1	DAF2_MOUSE	COMPLEMENT	DECAY-ACCEL	8.56e-35
39	232	11.4	372	1	LEM1_RAT	L-SELECTIN	PRECURSOR (2.87e-33
40	231	11.4	372	1	LEM1_MOUSE	L-SELECTIN	PRECURSOR (5.15e-33
41	219	10.8	330	1	CFH1_HUMAN	COMPLEMENT	FACTOR H-LI	5.42e-33
42	210	10.8	372	1	LEM1_MACMU	L-SELECTIN	PRECURSOR (3.04e-30
43	220	10.8	372	1	LEM1_PATRA	L-SELECTIN	PRECURSOR (3.04e-30
44	220	10.8	372	1	LEM1_PATHA	L-SELECTIN	PRECURSOR (3.04e-30
45	220	10.8	372	1	LEM1_HUMAN	L-SELECTIN	PRECURSOR (3.04e-30

ALIGNMENTS

ID	RESULT	1	STANDARD:	PRT:	1231 AA.
AC	PFAM_HUMAN				
AC	P08603:				
DT	01-AUG-1988 (Rel. 08, Created)				
DT	01-JAN-1990 (Rel. 13, Last sequence update)				
DT	15-FEB-2000 (Rel. 39, Last annotation update)				
DE	COMPLEMENT FACTOR H PRECURSOR.				
GN	HFI OR HF OR CFH.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
OC	Euteleia; Primates; Catarrhini; Homnidae; Homo.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=LIVER:				
RA	MEDLINE: 88134059.				
RA	Ripoche J., Day A.J., Harris T.J.R., Sim R.B.:				
RT	"The complete amino acid sequence of human complement factor H.":				
RL	Biochem. J. 249:593-602(1988).				
RN	[2]				
RN	SEQUENCE OF 53-445 FROM N.A.				
RX	MEDLINE: 87054207.				
RA	Schulz T.F., Schaeble W., Stanley K.K., Weiss E., Dierich M.P.:				
RT	"Human complement factor H: isolation of cDNA clones and partial cDNA				
RT	sequence of the 38-kDa tryptic fragment containing the binding site				
RT	for C3b.";				
RL	Eur. J. Immunol. 16:1351-1355(1986).				
RN	[3]				
RN	SEQUENCE OF 226-445 FROM N.A., AND PARTIAL SEQUENCE.				
RX	MEDLINE: 86169701.				
RA	Kristensen T., Wetsel R.A., Tack B.F.:				
RT	"Structural analysis of human complement protein H: homology with C4B				
RT	binding protein, beta2-glycoprotein I, and the Ba fragment of B2.";				
RL	J. Immunol. 136:3407-3411(1986).				
RN	[4]				
RN	SEQUENCE OF 1047-1231 FROM N.A.				
RP	MEDLINE: 91201892.				
RA	Esteller C., Koistinen V., Schaeble W., Dierich M.P., Weiss E.H.:				
RT	"Cloning of the 1.4-kb mRNA species of human complement factor H				
RT	reveals a novel member of the short consensus repeat family related				
RT	to the carboxy terminal of the classical 150-kDa molecule.";				
RL	J. Immunol. 146:3190-3196(1991).				
RN	[5]				
RP	SEQUENCE OF 19-35.				
RX	MEDLINE: 83048213.				
RA	Sim R.B., Discipio R.G.:				
RT	"Purification and structural studies on the complement-system control				
RT	protein beta 1H (Factor H).";				

Best Local Similarity 100.0%; Pred. No. 0.00e+00; Mismatches 265; Conservative 0; Indels 0; Gaps 0;

DB 19 EDCELPPEPNTTELLTSSWSQQTPEPTQAIYKCRPGYRISGNVIMCRGEMWALPLR 78
 OY 1 EDCELPPEPNTTELLTSSWSQQTPEPTQAIYKCRPGYRISGNVIMCRGEMWALPLR 60
 DB 79 KCOKRCPCGHPDPEFGFTLLTGGVNFEGYKAVYTCNEGVLLEINVRCECDTGWNTDI 138
 OY 61 KCOKRCPCGHPDPEFGFTLLTGGVNFEGYKAVYTCNEGVLLEINVRCECDTGWNTDI 120
 DB 139 PICEVVKCLPYTAPENGKIVSSAMEPDRYHFGQAVRFVCSGKIKGDEEMHCSDDGFW 198
 OY 121 PICEVVKCLPYTAPENGKIVSSAMEPDRYHFGQAVRFVCSGKIKGDEEMHCSDDGFW 180
 DB 199 SKERKCVETISCKSPDVINGSPISOKIYKENERFOYCNMGYSESGDGVCTESGMRP 258
 OY 181 SKERKCVETISCKSPDVINGSPISOKIYKENERFOYCNMGYSESGDGVCTESGMRP 240
 DB 259 LPSCCEKSCDNPYIPNGDYSPLRK 283
 OY 241 LPSCCEKSCDNPYIPNGDYSPLRK 265

RESULT 2
 ID CFAH MOUSE STANDARD: PRT: 1234 AA.
 AC P06909:
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE COMPLEMENT FACTOR H PRECURSOR (PROTEIN BETA-1-H).
 GN CFH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN (1)
 RN MEDLINE: 86233353.
 RA Kristensen T., Tack B.F.;
 RT "Murine protein H is comprised of 20 repeating units, 61 amino acids
 in length.";
 RT Proc. Natl. Acad. Sci. U.S.A. 83:3963-3967(1986).
 RN (2)
 RN SEQUENCE OF 1-19 FROM N.A.
 RC STRAIN-BALB/C;
 RX MEDLINE: 90148935.
 RA Munoz-Canoves P., Tack B.F., Vik D.P.;
 RT "Analysis of complement factor H mRNA expression: dexamethasone and
 IFN-gamma increase the level of H in L cells.";
 RT Biochemistry 28:9891-9897(1989).
 RN (3)
 RN SEQUENCE OF 1-18 FROM N.A.
 RX MEDLINE: 90111033.
 RA Natsune-Sakai S., Nonaka M., Nonaka M., Harada Y.N., Shreffler D.C.,
 RT "Demonstration of an unusual allelic variation of mouse factor H by
 the complete cDNA sequence of the H.2 allotype.";
 RT J. Immunol. 144:358-362(1990).
 CC -1- FUNCTION: FACTOR H FUNCTIONS AS A COFACTOR IN THE INACTIVATION OF
 C3B BY FACTOR I AND ALSO INCREASES THE RATE OF DISSOCIATION OF THE
 C3BB COMPLEX (C3 CONVERTASE) AND THE (C3)BB COMPLEX (C5
 CONVERTASE) IN THE ALTERNATIVE COMPLEMENT PATHWAY.
 CC -1- POLYMORPHISM: TWO CODOMINANT ALLELES OF FACTOR H ARE PRESENT IN
 MICE.
 CC -1- SIMILARITY: CONTAINS 20 SUSHI (SCR) REPEATS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: M12660; AAA37759.1; -
 DR EMBL: J02891; AAA37795.1; -
 DR EMBL: M31979; AAA37762.1; -
 DR PIR: A26154; NBM5H.
 DR HSP: P08603; IHFI.
 DR MGD; MGI:88385; CFH.
 DR PFAM; PF00084; sushi; 20.
 DR Complement alternate pathway; Plasma; Glycoprotein; Repeat; Sushi;
 KW Signal.
 FT SIGNAL.
 FT CHAIN 1 18
 FT DOMAIN 19 1234
 FT REPEAT 20 1224
 FT REPEAT 20 81
 FT REPEAT 84 142
 FT REPEAT 145 206
 FT REPEAT 209 263
 FT REPEAT 266 321
 FT REPEAT 324 386
 FT REPEAT 388 443
 FT REPEAT 447 506
 FT REPEAT 508 565
 FT REPEAT 568 623
 FT REPEAT 628 684
 FT REPEAT 689 744
 FT REPEAT 751 803
 FT REPEAT 807 862
 FT REPEAT 866 932
 FT REPEAT 935 990
 FT REPEAT 993 1049
 FT REPEAT 1052 1108
 FT REPEAT 1113 1169
 FT REPEAT 1171 1234
 FT DISULFID 21 66
 FT DISULFID 52 80
 FT DISULFID 85 129
 FT DISULFID 114 141
 FT DISULFID 146 192
 FT DISULFID 178 205
 FT DISULFID 210 251
 FT DISULFID 237 262
 FT DISULFID 267 309
 FT DISULFID 294 320
 FT DISULFID 325 374
 FT DISULFID 357 385
 FT DISULFID 389 431
 FT DISULFID 416 442
 FT DISULFID 448 494
 FT DISULFID 477 505
 FT DISULFID 509 553
 FT DISULFID 536 564
 FT DISULFID 569 610
 FT DISULFID 597 622
 FT DISULFID 629 672
 FT DISULFID 658 683
 FT DISULFID 690 732
 FT DISULFID 718 743
 FT DISULFID 752 791
 FT DISULFID 780 802
 FT DISULFID 808 850
 FT DISULFID 836 861
 FT DISULFID 867 920
 FT DISULFID 906 931
 FT DISULFID 936 978
 FT DISULFID 964 989
 FT DISULFID 994 1037
 FT DISULFID 1023 1048
 FT DISULFID 1053 1096
 FT DISULFID 1082 1107
 FT DISULFID 1114 1157
 FT DISULFID 1143 1168
 FT DISULFID 1172 1223
 FT DISULFID 1206 1233

FT	DISULFID	454	480		BY SIMILARITY.
FT	DISULFID	484	525		BY SIMILARITY.
FT	DISULFID	511	538		BY SIMILARITY.
FT	DISULFID	546	546		INTERCHAIN (WITH BETA CHAIN)
FT	DISULFID	558	558		INTERCHAIN (WITH BETA CHAIN) (POTENTIAL).
FT	CARBOHYD	221	221		
FT	CARBOHYD	506	506		
FT	CARBOHYD	528	528		
FT	VARIANT	92	92		
FT				Q -> T.	
FT				X -> H.	
FT	VARIANT	357	357	/FTId=VAR_001978.	
FT				/FTId=VAR_001978.	
SQ	SEQUENCE	597 AA;	67033 MW;	67E03FEA85A16DD CRC64;	
	Query Match	18.8%;	Score 382;	DB 1;	Length 597;
	Best Local Similarity	31.5%;	Pred. No. 3.17e-73;		
	Matches	79;	Conservative	52;	Mismatches 98; Indels 22; Gaps 20;
Dd	67	TETPRKKTGTLTKYLTCLPGVRSHTSQTLLCNDSGEV-VNTSTCIY-KR-CRHSGELRNQ	123	:	: : : : :
Oy	20	SDQRYPEGTQAIVKRCRPY-RSLGVNIWCKR-GERVALNPLRKCKRCRGHGDPFFET	77	:	: : : : :
Dd	124	VEIT-DL-STGSQLEBSCBSGFLLIGSTTSCEVDGRGVMSHPLOCEIYCKPPDI	181	:	: : : : :
Oy	78	FTLLGNVFEEGYNAVYTCTNEGQQLLGEINFR-EC-DTD-GWTNDIPICEVVKCLPYTAP	134	:	: : : : :
Dd	182	RNGR-HSG-E-ENFYAAFGSVYSGCCDRFSLLGHASISCTYENETIGVWRSPPTCEKI	237	:	: : : : :
Oy	135	ENGRIYSSAMPDDEYHNGQAVRFVCNSGKYKEGDENHC-S-D-D--GFWSREKPCACVI	190	:	: : : : :
Dd	238	TCKRPDYSHGEMVSFGPIYNYKDTIVKCOGFVLRGSSVYICDCADSKWNPSPPACEPN	297	:	: : : : :
Oy	191	SCKSPDIVINGSPISQ-KIIYKENERFOYCKMGMGYLSRGGAVC-TESGMR-LPSCCEK	247	:	: : : : :
Dd	298	SCNLPLPIDIPA	308		
Oy	248	SCDN-PYIPNG	257		
RESULT	5				
ID	APOH_MOUSE	STANDARD;	PRT;	345 AA.	
AC	001339;				
DT	01-APR-1993 (Rel. 25, Created)				
DT	01-APR-1993 (Rel. 25, Last sequence update)				
DT	15-JUL-1999 (Rel. 38, Last annotation update)				
DE	BETA-2-GLYCOPROTEIN I PRECURSOR (APOLIPOPROTEIN H) (APO-H) (ACTIVATED PROTEIN C-BINDING PROTEIN) (APC INHIBITOR) (BZGP1).				
GN	APOH OR BZGP1.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;				
CC	Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 92337200.				
RA	Nonaka M., Matsuda Y., Shirosaki T., Moriwa K., Natsume-Sakai S.;				
RL	"Molecular cloning of mouse beta 2-glycoprotein I and mapping of the gene to chromosome 11."				
RL	Genomics 13:1082-1087(1992).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CBA/J; TISSUE=LIVER;				
RX	MEDLINE; 94242017.				
RA	Sellar G.C., Steel D.M., Zafirooulos A., Seery L.T.,				
RT	Whitehead A.S.;				
RT	"Characterization, expression and evolution of mouse beta 2-				
RT	glycoprotein I (apolipoprotein H)."				
RL	Biochem. Biophys. Res. Commun. 200:1521-1528(1994).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=BALB/C; TISSUE=LIVER;				
RA	Kristensen T.;				
RT	"Structure of the human beta-2-glycoprotein I gene."				
RT	Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases				

```
CC -I- FUNCTION. BINDS TO VARIOUS KINDS OF NEGATIVELY CHARGED SUBSTANCES  
CC SUCH AS HEPARIN, PHOSPHOLIPIDS, AND DEXTRAN SULFATE. MAY PREVENT  
CC ACTIVATION OF THE INTRINSIC BLOOD COAGULATION CASCADE BY BINDING  
CC TO PHOSPHOLIPIDS ON THE SURFACE OF DAMAGED CELLS.  
CC -I- TISSUE SPECIFICITY: PLASMA.  
CC -I- SIMILARITY: CONTAINS 4 SUSHI (SCR) REPEATS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sdb.ch/  
CC or send an email to license@sdb.slb.ch).  
CC -----  
DR EMBL: D10096; BAA0094.5.1; -  
DR EMBL: S70439; AAB30789.1; -  
DR EMBL: Y11356; CAA72190.1; -  
DR PIR: A43286; NBSW.  
DR HSSP: P10998; IVOG.  
DR MGD: MGJ:88058; AVOH.  
DR PFAM: PF00084; sushi; 4.  
KW Heparin-binding; Glycoprotein; Plasma; Repeat; Sushi; Signal.  
FT FT SIGNAL  
FT CHAIN 1  
FT DOMAIN 20 345 BETA-2-GLYCOPROTEIN I.  
FT REPEAT 22 261 4 x SUSHI (SCR) REPEATS.  
FT REPEAT 22 80 SUSHI 1.  
FT REPEAT 83 138 SUSHI 2.  
FT REPEAT 141 201 SUSHI 3.  
FT REPEAT 204 261 SUSHI 4.  
FT DOMAIN 263 345 MODIFIED-SUSHI.  
FT DISULFD 23 66 BY SIMILARITY.  
FT DISULFD 51 79 BY SIMILARITY.  
FT DISULFD 84 124 BY SIMILARITY.  
FT DISULFD 110 137 BY SIMILARITY.  
FT DISULFD 142 188 BY SIMILARITY.  
FT DISULFD 174 200 BY SIMILARITY.  
FT DISULFD 205 248 BY SIMILARITY.  
FT DISULFD 234 260 BY SIMILARITY.  
FT DISULFD 264 315 BY SIMILARITY.  
FT DISULFD 300 325 BY SIMILARITY.  
FT CAROHDYD 307 345 BY SIMILARITY.  
FT CAROHDYD 162 182 BY SIMILARITY.  
FT CAROHDYD 183 193 POTENTIAL.  
FT CONFLICT 193 252 POTENTIAL.  
SQ SEQUENCE 345 AA; 38619 MW; C63FB6BBD51C940 CRC64;  
  
Query Match 18.3%; Score 371; DB 1; Length 345;  
Best Local Similarity 29.4%; Pred.No.3.37e-70;  
Matches 74; Conservative 49; Mismatches 110; Indels 19; Gaps 15;
```


DE COMPLEMENT CONTROL PROTEIN HOMOLOG PRECURSOR (CCPH).
 GN 4 OR CCPH.
 OS Herpesvirus saimiri (Strain 11).
 CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 9233368.
 RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,
 RA Newman C., Wiltman S., Craxton M.A., Coleman H., Fleckenstein B.,
 RA Honess R.W.;
 RT "Primary structure of the herpesvirus saimiri genome."
 RL J. Virol. 66:5047-5058(1992).
 RN [2]
 RP SIMILARITY TO CCP.
 RX MEDLINE; 92260674.
 RA Albrecht J.-C., Fleckenstein B.;
 RT "New member of the multigene family of complement control proteins in
 RT herpesvirus saimiri."
 RL J. Virol. 66:3937-3940(1992).
 CC -1- ALTERNATIVE PRODUCTS: A MEMBRANE-BOUND FORM AND A SECRETED FORM
 CC ARE PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE.
 CC -1- SIMILARITY: BELONGS TO THE SUPERFAMILY OF THE REGULATORS OF
 CC COMPLEMENT ACTIVATION (RCA).
 CC -1- SIMILARITY: CONTAINS 3 SUSHI (SCR) REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X64346; CAA45626.1; -
 CC EMBL; X64346; CAA43627.1; -
 CC DR EMBL; X60283; CAA42823.1; -
 CC DR EMBL; X60283; CAA42822.1; -
 CC DR PIR; B42534; WMBE2E.
 CC DR PIR; A42534; WMBE2E.
 CC DR PIR; S24567; S24567.
 CC HSSP; P10998; IYVC.
 CC DR PFM; PF00084; sush1; 4.
 CC KW Signal; Repeat; Sush1; Transmembrane; Alternative splicing;
 CC Glycoprotein.
 CC FT SIGNAL 1 20 POTENTIAL.
 CC FT CHAIN 21 360 COMPLEMENT CONTROL PROTEIN HOMOLOG.
 CC FT DOMAIN 83 265 3 X SUSHI (SCR) REPEATS.
 CC FT REPEAT 83 143 SUSHI 1.
 CC FT REPEAT 146 206 SUSHI 2.
 CC FT REPEAT 209 265 SUSHI 3.
 CC FT TRANSMEM 328 350 POTENTIAL.
 CC FT DISULFID 84 142 BY SIMILARITY.
 CC FT DISULFID 111 142 BY SIMILARITY.
 CC FT DISULFID 147 191 BY SIMILARITY.
 CC FT DISULFID 175 205 BY SIMILARITY.
 CC FT DISULFID 210 252 BY SIMILARITY.
 CC FT DISULFID 238 264 BY SIMILARITY.
 CC FT CARBOHYD 36 36 POTENTIAL.
 CC FT CARBOHYD 39 39 POTENTIAL.
 CC FT CARBOHYD 46 46 POTENTIAL.
 CC FT CARBOHYD 72 72 POTENTIAL.
 CC FT CARBOHYD 155 155 POTENTIAL.
 CC FT CARBOHYD 294 294 POTENTIAL.
 CC FT VARSPLIC 289 302 RINGNCTSMPTQ -> AECACPGSNYPIS (IN
 CC FT VARSPLIC 303 360 MISSING (IN SHORT ISOFORM).
 CC FT SEQUENCE 360 AA: 40006 MW: 6278A6C2EDD49669 CR664;
 CC
 CC Query Match 16.8%; Score 341; DB 1; Length 360;
 CC Best Local Similarity 31.9%; Pred. No. 5,236-62;
 CC Matches 74; Conservative 45; Mismatches 91; Indels 22; Gaps 17;

DB 44 YPNGTTLHVTGREGYAKRPVQITCVNGM-TV-P-KKQKKKSTPODLLNGRYITVG- 99
 OY 24 YPEGTQAIYKCPGGRYSIGNVIMCRKGEWALNPLRCKQRPGRGHPDPFFFTLTGG 83
 DB 100 NIX-YGSVITTYCNCSYLLISTT-SACLTRGGGVDMTPPPICDIKKRP--PQIAT 155
 OY 84 NVEYGVAVVYCNNGYDGLGINTREC--DTDG---WTNDIPICEYVKCLPYAPANGK 138
 DB 156 GHTTWK-DE-YTYIDTVYVSCNDETKLTLTPSSKLCSETGSWVNGETCEFIICKLP 213
 OY 139 IYSAMEPREDREHFGQAVRVCNSGYKIE--GDEMHCSGDGFV-SKEPKVCVEISCKSP 195
 DB 214 QVANAYVEVRKATSMQYLHIWVKCYKGMALYGEIPNT-CHNGWSPALPPC 264
 OY 196 DVING-SPISOKIYKENERFOYKCMNGE-YSERGDVCTESGWRP-LPSC 244
 RESULT 9
 ID C4BP.MOUSE STANDARD; PRT; 469 AA.
 AC P08607;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE C4B-BINDING PROTEIN PRECURSOR (C4BP).
 GN C4BPA OR C4BP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 88024997.
 RA Kristensen T., Ogata R.T., Chung L.P., Reid K.B.M., Tack B.F.;
 RT "CDNA structure of murine C4b-binding protein, a regulatory component
 RT of the serum complement system."
 RL Biochemistry 26:4668-4674(1987).
 CC -1- FUNCTION: C4BP CONTROLS THE CLASSICAL PATHWAY OF COMPLEMENT
 CC ACTIVATION. IT BINDS AS A COFACTOR TO C3B/C4B INACTIVATOR
 CC (C3BNA) WHICH THEN HYDROLYZES THE COMPLEMENT COMPLEX (C3
 CC ALSO ACCELERATES THE DEGRADATION OF THE COMPLEMENT FRAGMENT C2A. ALPHA
 CC CONVERTASE) BY DISSOCIATING THE COMPLEMENT COMPLEX (C3
 CC CHAIN BINDS C4B. IT INTERACTS ALSO WITH SERUM AMYLOID P COMPONENT.
 CC -1- SUBUNIT: HOMOPOLYMER; NOT COVALENTLY LINKED. MOUSE LACKS THE
 CC BETA CHAIN OF C4BP.
 CC -1- SIMILARITY: TO C4BP BETA CHAIN AND TO PIG APOLIPOPROTEIN R.
 CC -1- SIMILARITY: CONTAINS 6 SUSHI (SCR) REPEATS.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-44 IS THE INITIATOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M17122; AAA37312.1; ALT_INIT.
 CC DR PIR; A27117; NBMSC4.
 CC DR HSSP; P10998; IYVC.
 CC DR MGD; MGI:88229; C4BP.
 CC DR PFM; PF00084; sush1; 6.
 CC KW Complement pathway; Plasma; Glycoprotein; Repeat; Sush1; Signal.
 CC FT SIGNAL 1 56
 CC FT CHAIN 57 469 C4B-BINDING PROTEIN.
 CC FT DOMAIN 57 414 6 X SUSHI (SCR) REPEATS.
 CC FT REPEAT 57 116 SUSHI 1.
 CC FT REPEAT 119 177 SUSHI 2.
 CC FT REPEAT 180 241 SUSHI 3.
 CC FT REPEAT 244 300 SUSHI 4.
 CC FT REPEAT 302 356 SUSHI 5.
 CC FT REPEAT 358 414 SUSHI 6.
 CC FT DISULFID 58 103 BY SIMILARITY.
 CC FT DISULFID 88 115 BY SIMILARITY.
 CC FT DISULFID 120 160 BY SIMILARITY.

FT DISULFID 146 176 BY SIMILARITY.
 FT DISULFID 181 223 BY SIMILARITY.
 FT DISULFID 209 240 BY SIMILARITY.
 FT DISULFID 245 287 BY SIMILARITY.
 FT DISULFID 273 299 BY SIMILARITY.
 FT DISULFID 303 343 BY SIMILARITY.
 FT DISULFID 329 353 BY SIMILARITY.
 FT DISULFID 359 400 BY SIMILARITY.
 FT DISULFID 386 413 BY SIMILARITY.
 FT CARBOHYD 227 274 POTENTIAL.
 FT CARBOHYD 275 292 POTENTIAL.
 FT CARBOHYD 292 292 POTENTIAL.
 FT CARBOHYD 366 366 POTENTIAL.
 FT CARBOHYD 381 381 POTENTIAL.
 FT CARBOHYD 428 428 POTENTIAL.
 SQ SEQUENCE 469 AA; 51551 MW; 41E137CB8DC6321 CRC64;

Query Match 16.8%; Score 341; DB 1; Length 469;
 Best Local Similarity 28.9%; Pred. No. 5,23e-62;
 Matches 78; Conservative 57; Mismatches 109; Indels 26; Gaps 23;

DB 60 PPAIPALPASDVNRTPFSHTLKYECLEPGYGRGSRMMVYKPGSEW-ETIS-VS-CA 116
 QY 7 PPRNTEILGMSNDQ-Y-PEGTOAIKCRPGY-RSIGNVIMCR-KGEWVALNPLKQ 63
 DB 117 KKCRNPGYLDNG-Y-VNGETI-TFGSOLFSCQEGFLLVSSST-SSCEVRGKGVAMSNP 172
 QY 64 KRCHGHDGDPFTFTLTLGNVFEYGVKAVYTCNEGYYQLGELIYRCDI-D-G--WTND 119
 DB 173 FPCVIVKCGPPDISNGK-HSGF-E-DF-YRNHGSYTCDDGFRVLSGPFICGVYVK 228
 QY 120 IPICEVVKCLPVAPENGKIVSSAMEPDRYHFGQAVFCNSGKYLEGDEMHCS--DD 177
 DB 229 TVPVWSSPPTCKEIKISOPNHLGVIVSGKATYTHRDSVRLACINGTVLRGRHVLTCQ 288
 QY 178 GF--WEKREKCKVEISCKSPDIVINGSPISQ-KIIVKENEFQYKCMNGYYSERGDAVCT 234
 DB 289 GNGMSLSLPTCEP-DCDLPALVNGYYSM 317
 QY 235 ESG-WRPDPSCEKSCD-NPYIPNGDYSP 262

RESULT 10
 AC LEM3_RAT STANDARD; PRT: 768 AA.
 AC P98106;
 DT 01-FEB-1996 (Rel. 33 Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE P-SELECTIN PROTEIN (GRANULE MEMBRANE PROTEIN 140) (GMP-140) (PADGEM)
 DE (CD62P) (LEUKOCYTE-ENDOTHELIAL CELL ADHESION MOLECULE 3) (LECAM3).
 GN SELP.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 CC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG;
 RX MEDLINE: 94333817.
 RA Auchampach J.A., Oliver M.G., Anderson D.C., Manning A.M.;
 RT "Cloning, sequence comparison and in vivo expression of the gene
 RT encoding rat P-selectin.";
 RL Gene 145:251-255(1994).
 CC -I- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS
 CC TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
 CC INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
 CC LEUKOCYTES. THE LIGAND RECOGNIZED IS STALY-LWIS X.
 CC -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -I- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINED: SPLEEN,
 CC LUNG, BRAIN, LIVER, HEART, KIDNEY, THYMUS, SMALL INTESTINE.
 CC -I- INDUCTION: ACUTE INFLAMMATION (PROBABLY).
 CC -I- SIMILARITY: TO OTHER SELECTINS/LECTINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.

CC -I- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -I- SIMILARITY: CONTAINS 8 SUSHI (SCR) REPEATS; RAT P-LECTIN LACKS THE
 CC HUMAN SUSHI-2 EQUIVALENT.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L23088; AAA60325.1;
 CC HSSP: P16109; 1FSB.

DR PRINTS; PRO00343; SELECTIN.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS0186; EGF_2; 1.
 DR PROSITE; PS00615; C-TYPE LECTIN_1; 1.
 DR PROSITE; PS50041; C-TYPE LECTIN_2; 1.
 DR PFAM; PF00059; lectin_c7; 1.
 DR PFAM; PF00084; sushi; 8.
 DR
 DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
 DR Selectin; Signal; Sushi; Repeat.

KW
 FT SIGNAL 1 41
 FT CHAIN 42 768
 FT DOMAIN 42 709
 FT TRANSMEM 710 733
 FT DOMAIN 734 768
 FT DOMAIN 38 158
 FT DOMAIN 159 195
 FT DOMAIN 199 700
 FT REPEAT 199 258
 FT REPEAT 261 320
 FT REPEAT 323 382
 FT REPEAT 385 444
 FT REPEAT 447 506
 FT REPEAT 509 568
 FT REPEAT 579 638
 FT REPEAT 641 700
 FT DISULFID 60 158
 FT DISULFID 131 150
 FT DISULFID 168 183
 FT DISULFID 185 194
 FT DISULFID 200 244
 FT DISULFID 230 257
 FT DISULFID 262 306
 FT DISULFID 292 319
 FT DISULFID 324 368
 FT DISULFID 354 381
 FT DISULFID 386 430
 FT DISULFID 416 443
 FT DISULFID 448 492
 FT DISULFID 478 505
 FT DISULFID 510 554
 FT DISULFID 540 567
 FT DISULFID 580 624
 FT DISULFID 610 637
 FT DISULFID 642 686
 FT DISULFID 672 699
 FT CARBOHYD 45 45
 FT CARBOHYD 54 54
 FT CARBOHYD 107 107
 FT CARBOHYD 212 212
 FT CARBOHYD 347 347
 FT CARBOHYD 456 456
 FT CARBOHYD 603 603
 FT CARBOHYD 654 654
 FT CARBOHYD 661 661
 FT CARBOHYD 679 679
 FT SITE 756 759

Query Match 16.8%; Score 341; DB 1; Length 768;
 SQ SEQUENCE 768 AA; 83517 MW; 26FD7E8A5F3F1316 CRC64;
 FT SIGNAL 1 41
 FT CHAIN 42 768
 FT DOMAIN 42 709
 FT TRANSMEM 710 733
 FT DOMAIN 734 768
 FT DOMAIN 38 158
 FT DOMAIN 159 195
 FT DOMAIN 199 700
 FT REPEAT 199 258
 FT REPEAT 261 320
 FT REPEAT 323 382
 FT REPEAT 385 444
 FT REPEAT 447 506
 FT REPEAT 509 568
 FT REPEAT 579 638
 FT REPEAT 641 700
 FT DISULFID 60 158
 FT DISULFID 131 150
 FT DISULFID 168 183
 FT DISULFID 185 194
 FT DISULFID 200 244
 FT DISULFID 230 257
 FT DISULFID 262 306
 FT DISULFID 292 319
 FT DISULFID 324 368
 FT DISULFID 354 381
 FT DISULFID 386 430
 FT DISULFID 416 443
 FT DISULFID 448 492
 FT DISULFID 478 505
 FT DISULFID 510 554
 FT DISULFID 540 567
 FT DISULFID 580 624
 FT DISULFID 610 637
 FT DISULFID 642 686
 FT DISULFID 672 699
 FT CARBOHYD 45 45
 FT CARBOHYD 54 54
 FT CARBOHYD 107 107
 FT CARBOHYD 212 212
 FT CARBOHYD 347 347
 FT CARBOHYD 456 456
 FT CARBOHYD 603 603
 FT CARBOHYD 654 654
 FT CARBOHYD 661 661
 FT CARBOHYD 679 679
 FT SITE 756 759
 SQ SEQUENCE 768 AA; 83517 MW; 26FD7E8A5F3F1316 CRC64;
 FT SIGNAL 1 41
 FT CHAIN 42 768
 FT DOMAIN 42 709
 FT TRANSMEM 710 733
 FT DOMAIN 734 768
 FT DOMAIN 38 158
 FT DOMAIN 159 195
 FT DOMAIN 199 700
 FT REPEAT 199 258
 FT REPEAT 261 320
 FT REPEAT 323 382
 FT REPEAT 385 444
 FT REPEAT 447 506
 FT REPEAT 509 568
 FT REPEAT 579 638
 FT REPEAT 641 700
 FT DISULFID 60 158
 FT DISULFID 131 150
 FT DISULFID 168 183
 FT DISULFID 185 194
 FT DISULFID 200 244
 FT DISULFID 230 257
 FT DISULFID 262 306
 FT DISULFID 292 319
 FT DISULFID 324 368
 FT DISULFID 354 381
 FT DISULFID 386 430
 FT DISULFID 416 443
 FT DISULFID 448 492
 FT DISULFID 478 505
 FT DISULFID 510 554
 FT DISULFID 540 567
 FT DISULFID 580 624
 FT DISULFID 610 637
 FT DISULFID 642 686
 FT DISULFID 672 699
 FT CARBOHYD 45 45
 FT CARBOHYD 54 54
 FT CARBOHYD 107 107
 FT CARBOHYD 212 212
 FT CARBOHYD 347 347
 FT CARBOHYD 456 456
 FT CARBOHYD 603 603
 FT CARBOHYD 654 654
 FT CARBOHYD 661 661
 FT CARBOHYD 679 679
 FT SITE 756 759
 SQ SEQUENCE 768 AA; 83517 MW; 26FD7E8A5F3F1316 CRC64;

RT "Isolation from fetal bovine serum of an apolipoprotein-H-like
RT protein which inhibits thymidine incorporation in fetal calf
RT erythroid cells.";
RL Blochem. J. 267:261-264(1990).
CC -1- FUNCTION: BINDS TO VARIOUS KINDS OF NEGATIVELY CHARGED SUBSTANCES
CC SUCH AS HEPARIN, PHOSPHOLIPIDS, AND DEXTRAN SULFATE. MAY PREVENT
CC ACTIVATION OF THE INTRINSIC BLOOD COAGULATION CASCADE BY BINDING
CC TO PHOSPHOLIPIDS ON THE SURFACE OF DAMAGED CELLS.
CC
CC -1- TISSUE SPECIFICITY: PLASMA.
CC
CC -1- SIMILARITY: CONTAINS 4 SUSHI (SCR) REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.jsb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: I07303; AAA30382.1;
CC EMBL: X60065; CAA42669.1; -;
CC PIR: S23597; MBBO.
CC PIR: S09032; S09032.
CC PFAM: PF00084; sushi. 4.
CC Heparin-binding; Glycoprotein; Plasma; Repeat; Sushi; Signal.
CC SIGNAL 1
CC CHAIN 19
CC DOMAIN 20 345 BETA-2-GLYCOPROTEIN I.
CC REPEAT 22 261 4 X SUSHI (SCR) REPEATS.
CC REPEAT 22 80 SUSHI 1.
CC REPEAT 83 138 SUSHI 2.
CC REPEAT 141 201 SUSHI 3.
CC REPEAT 204 261 SUSHI 4.
CC DOMAIN 263 345 MODIFIED-SUSHI.
CC DISULFID 23 66
CC DISULFID 51 79
CC DISULFID 84 124
CC DISULFID 110 137
CC DISULFID 142 188
CC DISULFID 174 200
CC DISULFID 205 248
CC DISULFID 224 260
CC DISULFID 264 315
CC DISULFID 300 325
CC DISULFID 307 345
CC CARBOHYD 92 92
CC CARBOHYD 162 162
CC CARBOHYD 183 183
CC CARBOHYD 193 193
CC CARBOHYD 253 253
CC CARBOHYD 253 253
CC CONFLICT 101 101 E -> G (IN REF. 1).
CC CONFLICT 108 108 F -> S (IN REF. 1).
CC CONFLICT 177 177 H -> R (IN REF. 1).
CC CONFLICT 191 191 H -> N (IN REF. 3).
CC CONFLICT 194 194 W -> C (IN REF. 1).
CC CONFLICT 259 259 S -> N (IN REF. 1).
CC CONFLICT 302 302 K -> R (IN REF. 1).
CC CONFLICT 305 305 K -> R (IN REF. 1).
CC CONFLICT 329 329 H -> R (IN REF. 1).
CC CONFLICT 345 AA: 38252 MW: E117DAB609461C33 CRC64:
SO SEQUENCE

Query Match 16.58; Score 334; DB 1; Length 345;
Best Local Similarity 29.18; Pred. No. 4.13e-60;
Matches 69; Conservative 44; Mismatches 104; Indels 20; Gaps 16;

Db 39 RYEPGBOIYFSGVSGIRFTCLTGLM-PINTL-KCMPRVCPFAGILENGTVR 96
QY 22 QITPEGQAIYKCRPGTRSGNVM-VCR-KGEVVALNPKRQKRPCKGPGTPTT 79
Db 97 YT--TFEYNTISFSGHTGYLKASSAK-CTEEGKMSFDLVCAPITCPPIPKFAS 152
QY 80 LTGNGVEYGVKAVYTCNEGYYQLGEINRECDTG-WINDIDICEVVKCLPYTABENG 138
Db 153 LSYKPLAGNNSPFGSKAV-FKCLPHAMGNDIVTCTHGNT-QUPREVERCPFP 210

QY 139 I-VSSAMEPDRHYHGOAVRFVCSNGYKIEGDEMHCSDDGFWSKEXKVEISCK--S- 194
Db 211 PONGFNNHNPANVLYKKDPTFG--CHETYSLDGEPEVCSFGNWSAOPSCKA-SC 264
QY 195 PDV-INGSPISOKIITYKENRFQYCNMGYEISEGDVATCSG-WRPLPSCENSC 249

RESULT 14
ID MCP HUMAN STANDARD; PRT: 377 AA.
AC P15529;
DT 01-APR-1990 (rel. 14, Created)
DT 01-OCT-1996 (rel. 34, Last sequence update)
DT 15-FEB-2000 (rel. 39, Last annotation update)
DE MEMBRANE COFACTOR PROTEIN PRECURSOR (CD46 ANTIGEN) (TROPHOBLAST
DE LEUCOCYTE COMMON ANTIGEN) (TLX).
GN MCP.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 35-58.
RX MEDLINE: 88286080.
RA Lublin D.M., Liszewski M.K., Post T.W., Arce M.A., Le Beau M.M.,
RA Reberich M.B., Lemons R.S., Seya T., Atkinson J.P.,
RT Molecular cloning and chromosomal localization of human membrane
RT cofactor protein (MCP). Evidence for inclusion in the multigene
RL family of complement-regulatory proteins.";
RL J. Exp. Med. 168:181-194(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 93119658.
RC TISSUE-TESTIS;
RA Cervoni F., Fienichel P., Akhoundi C., Hsi B.L., Rossi B.,
RT Characterization of a cDNA clone coding for human testis membrane
RT cofactor protein (MCP, CD46)."
RL Mol. Reprod. Dev. 34:107-113(1993).
RN [3]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE: 91267562.
RA Purcell D.F., Russell S.M., Deacon N.J., Brown M.A., Hooker D.J.,
RA McKenzie I.F.;
RT "Alternatively spliced RNAs encode several isoforms of CD46 (MCP), a
RT regulator of complement activation.";
RL Immunogenetics 33:335-344(1991).
RN [4]
RP SEQUENCE OF 1-34 FROM N.A.
RX MEDLINE: 94014356.
RA Cui W., Hourcade D., Post T., Greenlund A.C., Atkinson J.P.,
RA Kumar V.;
RT "Characterization of the promoter region of the membrane cofactor
RT protein (CD46) gene of the human complement system and comparison to
RT a membrane cofactor protein-like genetic element.";
RL J. Immunol. 151:4137-4146(1993).
RN [5]
RP ALTERNATIVE SPLICING.
RX MEDLINE: 92289809.
RA Russell S.M., Sparrow R.L., McKenzie I.F.C., Purcell D.F.J.;
RT "Tissue-specific and allelic expression of the complement regulator
RT CD46 is controlled by alternative splicing.";
RL Eur. J. Immunol. 22:1513-1518(1992).
CC -1- FUNCTION: MCP IS AN IMPORTANT MEMBRANE PROTEIN FOR PROTECTING HOST
CC CELLS FROM DAMAGE BY COMPLEMENT. MCP POSSESSES COFACTOR ACTIVITY
CC FOR INACTIVATION OF C3B AND C4B BY SERUM FACTOR I. THIS WOULD
CC ALLOW MCP TO REGULATE THE COMPLEMENT SYSTEM AT THE LEVEL OF THE C3
CC CONVERTASES. MAY BE INVOLVED IN THE FUSION OF THE SPERMATOZOA WITH
CC THE OOCYTE DURING FERTILIZATION. MAY ACT BY PREVENTING COMPLEMENT
CC ACTIVATION AT THE FETO-MATERNAL INTERFACE ON THE
CC SYNCYTOTROPHOBLAST LAYER OF PLACENTA.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 14 FORMS ARE PRODUCED BY
CC ALTERNATIVE SPLICING. THE FORM SHOWN IS THE TRANSCRIPT A. THE
CC GAMMA ISOFORM (TRANSCRIPTS A AND B) IS PREFERENTIALLY EXPRESSED IN

RT alternation of anticardiolipin binding.";
 RL Int. Immunol. 3:1217-1221(1991).
 RP SEQUENCE FROM N.A.
 RA Kristensen T.;
 RT "Structure of the human beta-2-glycoprotein I gene."
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 20-345, CARBOHYDRATE-BINDING SITES, AND DISULFIDE BONDS.
 RX MEDLINE: 84222015.
 RA Lozier J., Takahashi N., Putnam F.W.;
 RT "Complete amino acid sequence of human plasma beta 2-glycoprotein I.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:3640-3644(1984).
 RN [8]
 RP DISULFIDE BONDS IN C-TERMINAL DOMAIN.
 RX MEDLINE: 93050249.
 RA Steinkasserer A., Barlow P.N., Willis A.C., Kertesz Z.,
 RT Campbell I.D., Sim R.B., Norman D.G.;
 RL "Activity, disulphide mapping and structural modelling of the fifth domain of human beta 2-glycoprotein I.";
 RN FEBS Lett. 313:193-197(1992).
 RL [9]
 RP ANALYSIS OF CARBOHYDRATES.
 RX MEDLINE: 97299942.
 RA Gambino R., Rulu G., Pagano G., Cassader M.;
 RT "Qualitative analysis of the carbohydrate composition of apolipoprotein H.";
 RL J. Protein Chem. 16:205-212(1997).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
 RX MEDLINE: 99437994.
 RA Bouma B., de Groot P.G., van Den Elsen J.M.H., Ravelli R.B.G.,
 RT Schouten A., Stimmelink M.J.A., Dersken R.H.W.M., Kroon J., Gros P.;
 RL "Adhesion mechanism of human beta(2)-glycoprotein I to phospholipids based on its crystal structure.";
 RN EMBO J. 18:5166-5174(1999).
 RL [11]
 RP VARIANT LEU-266.
 RX MEDLINE: 93273313.
 RA Steinkasserer A., Doerner C., Wuerzner R., Sim R.B.;
 RT "Human beta 2-glycoprotein I: molecular analysis of DNA and amino acid polymorphism.";
 RL Hum. Genet. 91:401-402(1993).
 RN [12]
 RP VARIANT ASN-107.
 RX MEDLINE: 97369481.
 RA Sanghera D.K., Kristensen T., Hamman R.F., Kamboh M.I.;
 RT "Molecular basis of the apolipoprotein H (beta 2-glycoprotein I) protein polymorphism.";
 RL Hum. Genet. 100:57-62(1997).
 RN [13]
 RP VARIANTS GLY-325 AND SER-335.
 RX MEDLINE: 9721791.
 RA Sanghera D.K., Wagenknecht D.R., McIntyre J.A., Kamboh M.I.;
 RT "Identification of structural mutations in the fifth domain of apolipoprotein H (beta-2-glycoprotein I) which affect phospholipid binding.";
 RL Hum. Mol. Genet. 6:311-316(1997).
 CC -I- FUNCTION: BINDS TO VARIOUS KINDS OF NEGATIVELY CHARGED SUBSTANCES SUCH AS HEPARIN, PHOSPHOLIPIDS, AND DEXTRAN SULFATE. MAY PREVENT ACTIVATION OF THE INTRINSIC BLOOD COAGULATION CASCADE BY BINDING TO PHOSPHOLIPIDS ON THE SURFACE OF DAMAGED CELLS.
 CC -I- TISSUE SPECIFICITY: PLASMA.
 CC -I- SIMILARITY: CONTAINS 4 SUSHI (SCR) REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: X58100; CAA4113.1; -
 DR EMBL: X53595; CAA37664.1; -
 DR EMBL: X57847; CAA40977.1; -
 DR EMBL: M62839; AAA51766.1; -
 DR EMBL: S80305; AAB21330.1; -
 DR EMBL: Y11493; CAA72279.1; -
 DR EMBL: Y11494; CAA72279.1; JOINED.
 DR EMBL: Y11495; CAA72279.1; JOINED.
 DR EMBL: X53595; CAA72279.1; JOINED.
 DR EMBL: Y11496; CAA72279.1; JOINED.
 DR EMBL: Y11497; CAA72279.1; JOINED.
 DR EMBL: Y11498; CAA72279.1; JOINED.
 DR PIR: S17178; NBRU.
 DR PIR: B43286; B43286.
 DR PDB: 1QUB; 08-OCT-99.
 DR MTM: 138700; -
 DR PIR: PFO0084; sushi; 4.
 DR Heparin-binding; Glycoprotein; Plasma; Repeat; Sushi; Signal;
 KW Polymorphism; 3D-structure.
 FT SIGNAL 1 19
 FT CHAIN 20 345
 FT DOMAIN 22 261
 FT REPEAT 22 80
 FT REPEAT 83 138
 FT REPEAT 141 201
 FT REPEAT 204 261
 FT DOMAIN 263 345
 FT DISULFID 23 66
 FT DISULFID 51 79
 FT DISULFID 84 124
 FT DISULFID 110 137
 FT DISULFID 142 188
 FT DISULFID 174 200
 FT DISULFID 205 248
 FT DISULFID 234 260
 FT DISULFID 264 315
 FT DISULFID 300 325
 FT DISULFID 307 345
 FT CARBOHYD 149 149
 FT CARBOHYD 162 162
 FT CARBOHYD 183 183
 FT CARBOHYD 193 193
 FT CARBOHYD 253 253
 FT VARIANT 107 107
 FT VARIANT 266 266
 FT VARIANT 325 325
 FT VARIANT 335 335
 FT CONFLICT 121 121
 FT CONFLICT 188 188
 FT SEQUENCE 345 AA; 38298 MW; 631017048DEFE3F CMC64;
 Query Match 16.4%; Score 333; DB 1; Length 345;
 Best Local Similarity 30.1%; Pred. No. 7.71e-60;
 Matches 72; Conservative 43; Mismatches 102; Indels 22; Gaps 18;
 Db 41 YERGERITTSCKRGYGRSGMRRKFCPLGLM-PINTL-KCPHYCPAGILNGAVRT 98
 Y 24 YPGTQAIYKCRPGYSLNVLN-VCR-KGEWVNLPLKCKRPGCGHPGDPFTFTLT 81
 Db 99 ---TFEPTNTISFCNTGFLNGADSAAK-CTEGCKWSPPLVPCAPILCPSPISPTFAFLR 154
 Y 82 GGNVFEIGVAVYTCNMGYGLGELIYVRCDDG-WTNDIPICEVYKCLPVPAPENKGI- 139
 Db 155 VYKPSAGNSNLYR-DTAV-PECLPQAHMGNDTIITTTGNNTKL-PEGREYKCFPSRP 211
 Y 140 V-SSAMEPDREYHFGQAVRVCSGYKIGDEMHCSDDGFWSKRPCKVEIISCKSPD-V 197

Fri Jun 9 10:53:38 2000

US-09-316-163-10.rsp

Page 16

Db 212 DNGFVNYPAKPTLYKKDKATFG--CHDGXSLDQPEIECTKLGNSAMPSCKA-SCKVP 267
OY 198 ING--S-PISOKIITYKENRFRQYKCNMGYEYSERGDVCTESG-WRPLPSCHEKSCDNP 252

Search completed: Thu Jun 8 21:42:12 2000
Job time : 13 secs.

WARNING

(TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MSrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jun 8 21:42:31 2000; Maspar time 26.07 Seconds
704.848 Million cell updates/sec

Tabular output not generated.

Title: >US-09-316-163-10
Description: (1-265) from US09316163.pep
Perfect Score: 2029
Sequence: 1 EDCNELPPRNTELLGWSWS.....EKSCDNPYPNGDYSPLRIK 265

Scoring table:
PAM 150
Gap 11

Searched: 225878 seqs, 69334122 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

sptrembl12
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human
5:sp-invertebrate 6:sp-mammal 7:sp-misc 8:sp-organelle
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified
13:sp-vertebrate 14:sp-virus

Statistics: Mean=43.031; Variance 62.408; scale 0.690

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	2029	100.0	449	4	Q14570	COMPLEMENT FACTOR H PR	0.00e+00
2	952	46.9	669	6	Q28085	CCP MODULES 3-12, WITH	4.00e-230
3	587	28.9	1053	13	Q91275	COMPLEMENT REGULATOR	7.83e-127
4	436	21.5	645	14	Q9WR02	COMPLEMENT BINDING PRO	2.63e-85
5	422	20.8	550	14	Q40912	ORF 04.	3.00e-81
6	421	20.7	550	14	P88903	ORF 4.	1.61e-81
7	398	19.6	259	14	P87616	41KBP FRAGMENT FROM LE	4.65e-75
8	396	19.5	657	4	Q14006	COMPLEMENT H FACTOR (F	1.60e-74
9	382	18.8	360	14	Q9Y708	COMPLEMENT CONTROL PRO	8.91e-71
10	357	17.6	263	14	Q89859	HOMOLOG OF VACCINIA VI	3.96e-64
11	357	17.6	263	14	Q07033	D12L PROTEIN.	3.96e-64
12	355	17.5	263	14	Q89076	B19L.	1.34e-63
13	353	17.4	263	14	Q22328	COSMID T07H6.	4.52e-63
14	351	17.3	533	11	Q08569	ACROSOMAL MATRIX COMPO	2.53e-62
15	350	17.2	263	14	Q89061	D15L.	2.81e-62
16	341	16.8	559	11	Q63135	COMPLEMENT REGULATOR	6.61e-60
17	338	16.7	349	4	Q15429	CD46.	4.07e-59
18	337	16.6	417	11	Q35520	512 ANTIGEN (FRAGMENT)	7.44e-59
19	336	16.6	740	4	Q95508	DJ780M13.1.2 (SELECTIN	1.36e-58
20	334	16.5	974	5	P91658	FURROWED.	4.56e-58

21	331	16.3	285	6	Q19126	MEMBRANE COFACTOR PROT	2.79e-57
22	331	16.3	285	6	Q19127	MEMBRANE COFACTOR PROT	2.79e-57
23	331	16.3	754	6	Q28290	CELL ADHESION MOLECULE	2.79e-57
24	328	16.2	369	6	P79138	MEMBRANE COFACTOR PROT	1.70e-56
25	328	16.2	2039	4	Q16745	COMPLEMENT RECEPTOR 1	1.70e-56
26	328	16.2	2489	4	Q16744	COMPLEMENT RECEPTOR 1	1.70e-56
27	326	16.1	579	11	Q60736	ZONA PELLICULA 3 RECP	5.68e-56
28	322	15.9	285	6	Q19121	MEMBRANE COFACTOR PROT	6.30e-55
29	322	15.9	285	6	Q19121	MEMBRANE COFACTOR PROT	6.30e-55
30	322	15.9	2014	6	Q29530	COMPLEMENT RECEPTOR 1	6.30e-55
31	320	15.8	522	6	Q28769	COMPLEMENT RECEPTOR (F	2.10e-54
32	318	15.7	377	6	Q62838	MEMBRANE COFACTOR PROT	6.96e-54
33	316	15.6	363	6	Q02839	MEMBRANE COFACTOR PROT	1.27e-53
34	317	15.6	378	6	Q02837	MEMBRANE COFACTOR PROT	1.27e-53
35	316	15.6	497	11	Q63612	512 ANTIGEN RECEPTOR	2.31e-53
36	313	15.4	1911	6	Q29528	COMPLEMENT RECEPTOR 1	1.38e-52
37	308	15.2	482	6	Q28982	E-SELECTIN.	2.76e-51
38	308	15.2	833	5	Q96442	FACTOR B SPBF.	2.76e-51
39	300	14.8	679	11	Q99254	COMPLEMENT RECEPTOR TY	3.24e-49
40	299	14.7	452	11	Q61407	COMPLEMENT FACTOR H-RE	5.88e-49
41	299	14.7	740	4	Q95507	COMPLEMENT FACTOR H-RE	5.88e-49
42	297	14.6	315	6	Q28770	DJ780M13.1.1 (SELECTIN	1.93e-48
43	295	14.5	347	4	P78361	COMPLEMENT RECEPTOR (F	6.31e-48
44	292	14.4	522	6	Q28797	DECAY-ACCELERATION FAC	3.72e-47
45	291	14.3	646	6	Q29097	UNKNOWN PROTEIN (FRAGM	6.74e-47
						P-SELECTIN PRECURSOR.	

ALIGNMENTS

RESULT 1
ID Q14570 PRELIMINARY; PRT; 449 AA.

AC Q14570; P78435; (TREMURel. 01, Created)
DT 01-NOV-1996 (TREMURel. 01, Last sequence update)
DT 01-NOV-1996 (TREMURel. 12, Last annotation update)
DE COMPLEMENT FACTOR H PRECURSOR.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.

RA R100CH J., DAY A.J., HARRIS T.J.R., SIM R.B.;
RX MEDLINE: 88134059.
RT "The complete amino acid sequence of human complement factor H.";
RN Biochem. J. 249:593-602(1988).

[2]
RN R100CH J., DAY A.J., HARRIS T.J.R., SIM R.B.;
RX MEDLINE: 86169701.
RT "Structural analysis of human complement protein H: homology with C4b
binding protein, beta2-glycoprotein I, and the Ba fragment of B2.";
RN J. Immunol. 136:3407-3411(1986).

[3]
RN R100CH J., DAY A.J., HARRIS T.J.R., SIM R.B.;
RX MEDLINE: 86169701.
RT "Structural analysis of human complement protein H: homology with C4b
binding protein, beta2-glycoprotein I, and the Ba fragment of B2.";
RN J. Immunol. 136:3407-3411(1986).

[4]
RN R100CH J., DAY A.J., HARRIS T.J.R., SIM R.B.;
RX MEDLINE: 86169701.
RT "Structural analysis of human complement protein H: homology with C4b
binding protein, beta2-glycoprotein I, and the Ba fragment of B2.";
RN J. Immunol. 136:3407-3411(1986).

[5]
RN R100CH J., DAY A.J., HARRIS T.J.R., SIM R.B.;
RX MEDLINE: 86169701.
RT "Structural analysis of human complement protein H: homology with C4b
binding protein, beta2-glycoprotein I, and the Ba fragment of B2.";
RN J. Immunol. 136:3407-3411(1986).

[6]
RN R100CH J., DAY A.J., HARRIS T.J.R., SIM R.B.;
RX MEDLINE: 86169701.
RT "Structural analysis of human complement protein H: homology with C4b
binding protein, beta2-glycoprotein I, and the Ba fragment of B2.";
RN J. Immunol. 136:3407-3411(1986).

[7]
RN R100CH J., DAY A.J., HARRIS T.J.R., SIM R.B.;
RX MEDLINE: 86169701.
RT "Structural analysis of human complement protein H: homology with C4b
binding protein, beta2-glycoprotein I, and the Ba fragment of B2.";
RN J. Immunol. 136:3407-3411(1986).

[8]
RN R100CH J., DAY A.J., HARRIS T.J.R., SIM R.B.;
RX MEDLINE: 86169701.
RT "Structural analysis of human complement protein H: homology with C4b
binding protein, beta2-glycoprotein I, and the Ba fragment of B2.";
RN J. Immunol. 136:3407-3411(1986).

[9]
RN R100CH J., DAY A.J., HARRIS T.J.R., SIM R.B.;
RX MEDLINE: 86169701.
RT "Structural analysis of human complement protein H: homology with C4b
binding protein, beta2-glycoprotein I, and the Ba fragment of B2.";
RN J. Immunol. 136:3407-3411(1986).

[10]
RN R100CH J., DAY A.J., HARRIS T.J.R., SIM R.B.;
RX MEDLINE: 86169701.
RT "Structural analysis of human complement protein H: homology with C4b
binding protein, beta2-glycoprotein I, and the Ba fragment of B2.";
RN J. Immunol. 136:3407-3411(1986).

[11]
RN R100CH J., DAY A.J., HARRIS T.J.R., SIM R.B.;
RX MEDLINE: 86169701.
RT "Structural analysis of human complement protein H: homology with C4b
binding protein, beta2-glycoprotein I, and the Ba fragment of B2.";
RN J. Immunol. 136:3407-3411(1986).

[12]
RN R100CH J., DAY A.J., HARRIS T.J.R., SIM R.B.;
RX MEDLINE: 86169701.
RT "Structural analysis of human complement protein H: homology with C4b
binding protein, beta2-glycoprotein I, and the Ba fragment of B2.";
RN J. Immunol. 136:3407-3411(1986).

[13]
RN R100CH J., DAY A.J., HARRIS T.J.R., SIM R.B.;
RX MEDLINE: 86169701.
RT "Structural analysis of human complement protein H: homology with C4b
binding protein, beta2-glycoprotein I, and the Ba fragment of B2.";
RN J. Immunol. 136:3407-3411(1986).

[14]
RN R100CH J., DAY A.J., HARRIS T.J.R., SIM R.B.;
RX MEDLINE: 86169701.
RT "Structural analysis of human complement protein H: homology with C4b
binding protein, beta2-glycoprotein I, and the Ba fragment of B2.";
RN J. Immunol. 136:3407-3411(1986).

[15]
RN R100CH J., DAY A.J., HARRIS T.J.R., SIM R.B.;
RX MEDLINE: 86169701.
RT "Structural analysis of human complement protein H: homology with C4b
binding protein, beta2-glycoprotein I, and the Ba fragment of B2.";
RN J. Immunol. 136:3407-3411(1986).

[16]
RN R100CH J., DAY A.J., HARRIS T.J.R., SIM R.B.;
RX MEDLINE: 86169701.
RT "Structural analysis of human complement protein H: homology with C4b
binding protein, beta2-glycoprotein I, and the Ba fragment of B2.";
RN J. Immunol. 136:3407-3411(1986).

[17]
RN R100CH J., DAY A.J., HARRIS T.J.R., SIM R.B.;
RX MEDLINE: 86169701.
RT "Structural analysis of human complement protein H: homology with C4b
binding protein, beta2-glycoprotein I, and the Ba fragment of B2.";
RN J. Immunol. 136:3407-3411(1986).

[18]
RN R100CH J., DAY A.J., HARRIS T.J.R., SIM R.B.;
RX MEDLINE: 86169701.
RT "Structural analysis of human complement protein H: homology with C4b
binding protein, beta2-glycoprotein I, and the Ba fragment of B2.";
RN J. Immunol. 136:3407-3411(1986).

[19]
RN R100CH J., DAY A.J., HARRIS T.J.R., SIM R.B.;
RX MEDLINE: 86169701.
RT "Structural analysis of human complement protein H: homology with C4b
binding protein, beta2-glycoprotein I, and the Ba fragment of B2.";
RN J. Immunol. 136:3407-3411(1986).

[20]
RN R100CH J., DAY A.J., HARRIS T.J.R., SIM R.B.;
RX MEDLINE: 86169701.
RT "Structural analysis of human complement protein H: homology with C4b
binding protein, beta2-glycoprotein I, and the Ba fragment of B2.";
RN J. Immunol. 136:3407-3411(1986).

Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 19 EDCELPFRATELTGSMQTYEGTQAIYKCPGRSGNVTMCRKGEWALNPLR 78
QY 1 EDCELPFRATELTGSMQTYEGTQAIYKCPGRSGNVTMCRKGEWALNPLR 60
Db 79 KCQRPCGHPDTEFGFTLLGNVFEYGVAVYTCNGYOLGELINRECDTGWTNDI 138
QY 61 KCQRPCGHPDTEFGFTLLGNVFEYGVAVYTCNGYOLGELINRECDTGWTNDI 120
Db 139 PICVAVCLPTAPENKIVSSAMEPDEYHFGQAVREVCNGYKIEGDEDMHCSDDGF 198
QY 121 PICVAVCLPTAPENKIVSSAMEPDEYHFGQAVREVCNGYKIEGDEDMHCSDDGF 180
Db 199 SKERPKCVESICKSPDINGSISQKIIYKNERFQYKCNMGYERSGDAVCTESGMR 258
QY 181 SKERPKCVESICKSPDINGSISQKIIYKNERFQYKCNMGYERSGDAVCTESGMR 240
Db 259 LPSCSEKSCNPIYIPNDYSPLRIT 283
QY 241 LPSCSEKSCNPIYIPNDYSPLRIT 265

RESULT 2
ID 028085 PRELIMINARY; PRT; 669 AA.
AC 028085;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, last annotation update)
DE CCP MODULES 3-12, WITH PARTS OF CCP 2 AND 13 (FRAGMENT).
OS BCS taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;
OC Bovinae; Bos.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE; 96202005.
RA SOAMES C.J., DAY A.J., SIM R.B.;
RT "Prediction from sequence comparisons of residues of factor H involved
in the interaction with complement component C3b."
RL Biochem. J. 315:523-531(1996).
DR HSSP; P10998; IYVC.
DR PFAM; PF00084; sushi; 11.
FT NON_TER 1
FT NON_TER 669
SQ SEQUENCE 669 AA; 75683 MW; FAF0D174 CRC32;

Query Match 46.9%; Score 952; DB 6; Length 669;
Best Local Similarity 64.3%; Pred. No. 4.00e-230;
Matches 117; Conservative 29; Mismatches 36; Indels 0; Gaps 0;

Db 7 LAEGQFEYGAQVYTCDEGTVMEANFRCDTNGTNDIPICVAVCLPYTEPENGKI 66
QY 80 LTGQVFEYGAQVYTCDEGTVMEANFRCDTNGTNDIPICVAVCLPYTEPENGKI 139
Db 67 FSDALEPQDEYTGQVYFECNSGYMLDGPQKHCSAGSWSAEKPKVEIFCKRPVILN 126
QY 140 VSSAMEPDEYHFGQAVREVCNGYKIEGDEDMHCSDDGF 199
Db 127 GQAVLPKATYQNEVQYRCAGFEYGGRGDTVCTKSGWTPAPTCIETCDPPRIPNGVY 186
QY 200 GSPISQKIYKNERFQYKCNMGYERSGDAVCTESGMRPLPSCSEKSCNPIYIPNDY 259
Db 187 RP 188
QY 260 SP 261

RESULT 3
ID 091275 PRELIMINARY; PRT; 1053 AA.

AC 091275;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, last annotation update)
DE COMPLEMENT REGULATORY PLASMA PROTEIN.
OS Paratubax nebulifer (barred sand bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
OC Perciformes; Percoidae; Serranidae; Paratubax.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE; 94318039.
RA DAMEY A., KAIHOD T., ZIPPEL P.F., GIGLI I.;
RT "Cloning and characterization of a cDNA representing a putative
complement-regulatory plasma protein from barred sand bass (Paratubax
nebulifer)."
RL Biochem. J. 301:391-397(1994).
DR EMBL; L21703; AAA82556.1; -.
DR HSSP; P08603; IHEH.
DR PFAM; PF00084; sushi; 16.
SQ SEQUENCE 1053 AA; 117597 MW; 0DF68EDB CRC32;

Query Match 28.9%; Score 587; DB 13; Length 1053;
Best Local Similarity 37.1%; Pred. No. 7.83e-127;
Matches 91; Conservative 41; Mismatches 101; Indels 12; Gaps 9;

Db 47 EASYEGRGVAVGCVNGY-S-GFKLVCEYGRWETRGA--KCOPRSCGHPDAOFADPHL 102
QY 21 DQYTPGTOAIYKCRFGYSLGNVIMCRKGEWALNPLRKCKCRGCHPGDTPETFTL 80
Db 103 AEGNDVFSGKYVYTCQKQYQVAVRYNRYCAEAGDGVVPCESQCC-PLIHVDNNVQV 161
QY 81 TGSNVFEYGVAVYTCNEGYOLGELINRECDTGWTNDIPICVAVCLPYAPENGKIV 140
Db 162 IGG--PE-EATFGNVVRFSCSRSEILDSPELCYDERGDSVPKCAICAIPIEN 218
QY 141 SSAMEPDEYHFGQAVREVCNGYKIEGDEDMHCSDDGF 199
Db 219 GNVPGAIRKYKNDVLAHECDRAFKHIDRPSTCIGIKRAEMSPPLIESIKRLLTMDG 278
QY 200 GSPISQKIYKNERFQYKCNMGYERSGDA-V--CTESGMRPLPSCSEKSCNPIYIPN 256
Db 279 TRYEP 283
QY 257 GDYSP 261

RESULT 4
ID 09WRU2 PRELIMINARY; PRT; 645 AA.
AC 09WRU2;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, last annotation update)
DE COMPLEMENT BINDING PROTEIN.
OS Macaca mulatta rhadinovirus 17577.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE; 99174001.
RA SEARLES R.P., BERGOUAM E.P., AXTHELM M.K., WONG S.W.;
RT "Sequence and genomic analysis of a Rhesus macaque rhadinovirus with
similarity to Kaposi's sarcoma-associated herpesvirus/human
herpesvirus 8."
RL J. Virol. 73:3040-3053(1999).
DR EMBL; AF083501; AAD21332.1; -.
SQ SEQUENCE 645 AA; 71526 MW; 60FB82D6 CRC32;

Query Match 21.5%; Score 436; DB 14; Length 645;
Best Local Similarity 30.9%; Pred. No. 2.63e-85;
Matches 72; Conservative 50; Mismatches 94; Indels 17; Gaps 14;

```

Db 337 EKVSVASVELLCRPGFTKMOSTVSECLSNGTWAPNA--KCHAKKCP7POELNGEYI 394
QY 22 QTPGEGQALYKCRPGRSIGNIWM-C-RKGEVALNPLKRCOKRPGCGHPTGPT 79
Db 395 VTSGEDAFKGTITTKCNNGCYOLLGSMVRIICMLKDLKVDDEPRAPICDICKCPPO 454
QY 80 LTGG-NVFEYGVAVYTCNEGQYLLG-E-IN-Y--REC-DTIDGNTINDIPICEVYKCLPYTA 133
Db 455 ITNGKY--HPEVK-DE-YQYLDVTFTSCNRDPSLVGDEMTTCISNT--WKPFRCEQITCS 509
QY 134 PENKTIYSAMEPREHFEQAVAFVCSNGYKIEGDEMHCSDDGEFWSKEKPCVETISCK 193
Db 510 APNIAHKKLLTGSSVYKRGOSTVIGETGFTLIGSEIISCKOSSMPPTC 562
QY 194 SPDVINGSPIS-QKIYKENERFOYKCMNGEYSEKGDVCTESGM-RPLPSC 244

RESULT 5
ID 040912 PRELIMINARY: PRT: 550 AA.
AC 040912:
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE ORF 04
OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97296220.
RA NEIRPEL F., ABBRECHT J.C., FLECKENSTEIN B.;
"Cell-homologous genes in the Kaposi's sarcoma-associated rhadinovirus
human herpesvirus 8: determinants of its pathogenicity?";
RL J. Virol. 71:4187-4192(1997).
DR EMBL: U93872; AAB62602.1; -.
DR HSSP: P10998; IYVC.
DR PFAM: PF00084; sush1.4.
SQ SEQUENCE 550 AA: 60648 MW: 558089B4 CRC32;

Query Match 20.8%; Score 422; DB 14; Length 550;
Best Local Similarity 31.9%; Pred. No. 1,61e-81;
Matches 75; Conservative 46; Mismatches 95; Indels 19; Gaps 15;

Db 52 RCRSGYTYARNITATCLOGGTW-S-EPTATCNKKSCEPNGEIONGKVIFFHGODALXYG 109
QY 33 KCRPGRSUG-NVIMYC-RKGEVALNPLKRCOKRPGCGHPTGPTLTLTG-NVFEYX 89
Db 110 ANISYVCNDEYFLVGREYRYCMIGASGOMANSSPPFCEKEK-H-R-P--KIKNDF 163
QY 90 VKAVYTCNEGQYLLG-E-INRECDTDG--WTNDIPICEVYKCLPYTAPENGKIVSSAM 144
Db 164 KPDKDYEVNDVAHFECNNGCYTLVGHSTACAVANNWTSMPTCELACGKPSVTHGYPI 223
QY 145 EPREYH-FEQAVAFVCSNGYKIEGDEMHCSDDGEFWSKEKPCVETISCKSPVINGSPI 203
Db 224 QGSLATYKHKOSTYFACNDGFLVRSPTTTCNTVENDPPLPKVCLDIDDPNNNS 278
QY 204 SQ-KIYKENERFOYKCMNGEYSEKGDVCTESGM-RPLPSCSEKSCDNPYIPN 256

RESULT 6
ID P88903 PRELIMINARY: PRT: 550 AA.
AC P88903:
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE ORF 4.
OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97121460.

```

```

RA RUSSO J.J., BOHENZKY R.A., CHIEN M.C., CHEN J., YAN M., MADDALENA D.,
RA PARRY J.P., PERIZZI D., EDELMAN I.S., CHANG Y., MOORE P.S.;
"Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus
(HHV8).";
RL Proc. Natl. Acad. Sci. U.S.A. 93:14862-14867(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97094384.
RA MOORE P.S., BASHOFF C., WEISS R.A., CHANG Y.;
"molecular mimicry of human cytokine and cytokine response pathway
genes by KSHV.";
RL Science 274:1739-1744(1996).
DR EMBL: U75698; AAC57082.1; -.
DR HSSP: P10998; IYVC.
DR PFAM: PF00084; sush1.4.
SQ SEQUENCE 550 AA: 60687 MW: CEA2CAEC CRC32;

Query Match 20.7%; Score 421; DB 14; Length 550;
Best Local Similarity 31.9%; Pred. No. 3.00e-81;
Matches 75; Conservative 46; Mismatches 95; Indels 19; Gaps 15;

Db 52 RCRSGYTYARNITATCLOGGTW-S-EPTATCNKKSCEPNGEIONGKVIFFHGODALXYG 109
QY 33 KCRPGRSUG-NVIMYC-RKGEVALNPLKRCOKRPGCGHPTGPTLTLTG-NVFEYX 89
Db 110 ANISYVCNDEYFLVGREYRYCMIGASGOMANSSPPFCEKEK-H-R-P--KIKNDF 163
QY 90 VKAVYTCNEGQYLLG-E-INRECDTDG--WTNDIPICEVYKCLPYTAPENGKIVSSAM 144
Db 164 KPDKDYEVNDVAHFECNNGCYTLVGHSTACAVANNWTSMPTCELACGKPSVTHGYPI 223
QY 145 EPREYH-FEQAVAFVCSNGYKIEGDEMHCSDDGEFWSKEKPCVETISCKSPVINGSPI 203
Db 224 QGSLATYKHKOSTYFACNDGFLVRSPTTTCNTVENDPPLPKVCLDIDDPNNNS 278
QY 204 SQ-KIYKENERFOYKCMNGEYSEKGDVCTESGM-RPLPSCSEKSCDNPYIPN 256

RESULT 7
ID P87616 PRELIMINARY: PRT: 259 AA.
AC P87616:
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE 41KB FRAGMENT FROM LEFT END OF GENOME.
GN D17L OR C17L.
OS Cowpox virus (CPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-GRI-90;
RX MEDLINE: 97068532.
RA SAFRONOV P.F., PETROV N.A., RIAZANKINA O.I., TOTMENIN A.V.,
RA SHEHELKUNOV S.N., SANDAKHCHIEV L.S.;
"Genes of a circle of hosts for the cowpox virus.";
RL Dokl. Akad. Nauk 349:829-833(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-GRI-90;
RX MEDLINE: 98229462.
RA SHEHELKUNOV S.N., SAFRONOV P.F., TOTMENIN A.V., PETROV N.A.,
RA RIAZANKINA O.I., GUTOROV V.V., KOTVAL G.J.;
"the genomic sequence analysis of the left and right species-specific
terminal region of a cowpox virus strain reveals unique sequences and
a cluster of intact ORFs for immunomodulatory and host range
proteins.";
RL Virology 243:432-460(1998).
DR EMBL: X94355; CA64102.1; -.
DR HSSP: Y11842; CAA72567.1; -.
DR HSSP: P10998; IYVC.
DR PFAM: PF00084; sush1.4.
SQ SEQUENCE 259 AA: 28193 MW: 76531F63 CRC32;

```

Query Match 19.6%; Score 398; DB 14; Length 259;
 Best Local Similarity 34.3%; Pred. No. 4,65e-75;
 Matches 86; Conservative 39; Mismatches 104; Indels 22; Gaps 18;

DB 20 CCEPISRPITKFKGT-VDSHYNIGDTIEYLCLPGRKQKMGPIYACITGTG-TL-FN 75
 3 CNEPAPRNNEILGMSDSOTYPEGTOAIKCRPGYRS--LGNIWVCRGGEVWALNPLR 60
 76 QCIRRCRCPSPRIDNGQIDT-GG-V-DEGSSITYSCNSGYHLIGESKSCYCELGSTGSMW 132
 61 KQCRPCGHHGDDFFGFTLGGVNFYGVKAVYTCNEGVLLEIN-YRE-CDTGG--W 116
 133 NPEAPICSGCCSPSISNGR--HNGYE-D-FYTDGSVTYTSCNSGYSLGNSGYLCSG 188
 117 TNDIPICEVVKCLPVTAPENGKIVSSAMEPDRYHFGQAVRFVCSNGYKIEGDEMHCSG 176
 189 GE-WS-DEPTQIVKCPHTISNGYLSGFRGYSYNDVDFCKCYKLSGSSSTGSP 246
 177 DGFWSKPKPCVBISSKSPDVINGSPISQ-KITYKENERQYKCMNGYSESGDAVCTE 235
 247 GNMWELPRC 257
 236 SG-WRP-LBSC 244

RESULT 8
 ID 014006 PRELIMINARY: PRT: 657 AA.
 AC 014006:
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DE 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE COMPLEMENT H FACTOR (FRAGMENT).
 GN HF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 88025472.
 RA DAY A.J., RIBOCHE J., LYONS A., MCINTOSH B., HARRIS T.J., SIM R.B.;
 RT "Sequence analysis of a cDNA clone encoding the C-terminal end of
 RL Biosci. Rep. 7:201-207(1987).
 DR EMBL: M17517; AAA52016.1; -;
 DR HSSP: P08603; 1HFH.
 DR PFM: PFO0084; sushi; 11.
 DR NON TER
 SQ SEQUENCE 657 AA; 74247 MW; F4AB5238 CRC32;

Query Match 19.5%; Score 396; DB 4; Length 657;
 Best Local Similarity 27.5%; Pred. No. 1.60e-74;
 Matches 74; Conservative 64; Mismatches 111; Indels 20; Gaps 19;

DB 296 CSO-PPOIEHTISSRSQESYAHGKLTCTCGGRISENETTCYMGW-S-SP-PQ 351
 3 CNEPAPRNNEILGMSDSOTYPEGTOAIKCRPGYRSIGNVIMVCRGGEVWALNPLR 61
 352 CEGLPKSPPEISHGVVHMSD-YQYGEVYTKCFEGGIDEPALAK-CLGKRNHP-P 408
 62 CQKPCGHPDPTGFTTLGNGVNFYGVKAVYTCNEGVLLEINRECDTGMWINDIP 121
 409 SCIKTDLSPSPENA-IPMGEEK-D-VYKAGQVYTYTCATYKMDGASAVTICINSR-WT 464
 122 ICEVWCLPYTAENGKIVSSAMEPDRYHFGQAVRFVCSNGYKIEGDEMHCSGSDGFS 181
 465 -GRPTGRDTCVNPPTQVNAIVSRMSKPSSEKRYRCRSRYEMFGD-EEVYMLNGW 522
 182 KEKPKCEVISC-KSPDVINGSPIS-OKITKENERQYKCMNGYE-YSRBGDAVCTESGW 238
 523 TEPPQCKDSTGKCGPPPIDNGDITSEPL 551
 239 RPLPSCERKS--CDNPY-IPNGDYSPLRI 264

RESULT 9
 ID 09YNO8 PRELIMINARY: PRT: 360 AA.
 AC 09YNO8:
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DE 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE COMPLEMENT CONTROL PROTEIN HOMOLOG CCPH.
 OS Ateles herpesvirus 3.
 OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-73;
 RA ADRECHT J.-C., FLECKENSTEIN B.;
 RT "Primary structure of the Herpesvirus Ateles Genome";
 RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF083424; AAC95530.1; -;
 DR HSSP: P10998; 1VYC.
 SQ SEQUENCE 360 AA; 40208 MW; F66C7ADA CRC32;

Query Match 18.8%; Score 382; DB 14; Length 360;
 Best Local Similarity 34.7%; Pred. No. 8.91e-71;
 Matches 82; Conservative 39; Mismatches 95; Indels 20; Gaps 15;

DB 40 SSGSYRPGTTLQYTCRKGITGRQITVTVYVGNW-TV-P-NECQRRCSTPADLINGWT 96
 20 SDQTYPEGTOAIKCRPGYRSIGNVIMVCRGGEVWALNPLRCKQKPCGHPDPTGFTT 79
 97 VTG-NLY-YGSVITYTCNGYOLGSP-T-SSCLGPDGVMNTPRPICETIKCKAPPTI 153
 80 LTGAVNEFYGVKAVYTCNEGVLLEINREC-DTG--WTNDIPICEVVKCLPVAP 134
 154 ANGT--HTNK-E-YTYTDAVYTSNDETKLTLTGSSKQCSSETGRWVDEETKCEFKV 209
 135 ENKIVSSAMEPDRYHFGQAVRFVCSNGYKIE--GDEMHCSDDGFW-SKEKPCVBI 191
 210 CKIPQVANHVEVRKTSNNVQYQVYNICDKGRFRLQGETPNNCKAGVFPALPTCE 265
 192 CKSPDVINGSPISOKITYKENERF-QYKCMNGYSESGDAVCTESGWRP-LPSC 245

RESULT 10
 ID 089859 PRELIMINARY: PRT: 263 AA.
 AC 089859:
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DE 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE HOMOLOG OF VACCINIA VIRUS CDS C31.
 GN D15L.
 OS Variola virus.
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BANGLADESH-1975;
 RX MASSUNG R.F., ESPOSITO J.J., LIU L.I., QI J., UTTERACK T.R.,
 RA KNIGHT J.C., AUBIN L., YURAN T.E., PARSONS J.M., LOPAREV V.N.;
 RT "Potential virulence determinants in terminal regions of variola
 RL smallpox virus genome.";
 Nature 366:748-751(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SOMALIA-1977;
 RA MASSUNG R.F., LOPAREV V.N., KNIGHT J.C., CHIZHIKOV V.E., PARSONS J.M.,
 RA TOMMENIN A.V., SHEKELTONOV S.N., ESPOSITO J.J.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: L22579; AAA60760.1; -;
 DR EMBL: U18340; AAA69423.1; -;
 DR HSSP: P10998; 1VYC.
 DR PFM: PFO0084; sushi; 4.

Query Match	17.48;	Score 353;	DB 5;	Length 560;
Best Local Similarity	32.28;	Pred. NO. 4.52e-63;		
Matches 68;	Conservative	42;	Mismatches 82;	Indels 19;
				Gaps 15;

RESULT	14		
ID	008569	PRELIMINARY;	PRT; 533 AA
NC	008569.		

BT	01-JUL-1997	(TRIMBLrel. 04, Created)
DT	01-JUL-1997	(TRIMBLrel. 04, Last sequence update)
DE	01-NOV-1999	(TRIMBLrel. 12, Last annotation update)
DD	ACROSOMAL MATRIX COMPONENT AM67	PRCUCSOR.
DD	Cavia porcellus (Guinea pig).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;	
OC	Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia,	
OC	[1]	
RP	SEQUENCE FROM N.A.	
RP	STRAIN-HARTLEY; TISSUE-TESTIS;	
RC	MEDLINE: 97284752	
RC	FOSTER J.A., FRIDAY B.B., MAULIT M.T., BLOEBL C., WINFREY V.P.,	
RA	OLSON G.E., KIM R.S., GERTON G.L.;	
RA	"AM67, a secretory component of the guinea pig sperm acrosomal matrix,	
RT	is related to mouse sperm protein sp56 and the complement component 4-	
RT	binding proteins.";	
RT	J. Biol. Chem. 272:12714-12722(1997).	
RDR	EMBL: U75654; AAC13868.1; -.	
RDR	HSSP: P10938; IYVC.	
RDR	PFAM: PF00084; sush1; 7.	
RDR	Signal.	
KM	SIGNAL.	
ET	CHAIN	1 28
ET	POTENTIAL.	
ET	ACROSOMAL MATRIX COMPONENT AM67.	
ET	SEQUENCE	533 AA: 59772 MW; 60DEA526 CRC32;

RESULT	15	
ID	Q89061	PRELIMINARY;
AC	Q89061	PRT; 263 AA.

Best Local Similarity	33.68;	Pred. No. 2.81e-62;	Denclun 203;
Matches 79; Conservative		37; Mismatches 96;	Indels 23; Gaps 19;

Db 40 AANNVINGDTIELCLCPYGRKQMGVIYAKCTGTG-TL--FNOCIKRCPSPRIDNGH 96
 QY 20 SDQTEPGTGAIKCAPGRS--LGNVIMCKRGEVWALNPLRKQKRCRCPHGPOTPGT 77
 Db 97 LDI-GG-V-DFGSSITYSCNSGYLLIGEEKYSYKLGSTGSMYMNRPAPICESVRCQLPS 153
 QY 78 FTLLGNVEYEGYKAVAYTCNEGYQLLGEIN-Y-RECDTUG--WTNDIPICEVVC-LPVT 132
 Db 154 IS-NGR--HNGVN-DE-YTDSVYIYSCNSGSLIGNSVLCGGE-MSNP-PTCIQYKC 206
 QY 133 APENGRIVSSAEPDEHYHFGAVRVCNSGGKIGIDEMHCSDDGFWSKEKPCVETISC 192
 Db 207 PHEITINGLSLGSFKRSYSYNDNVPTCKCYGKLLGSLSLTSCSPNTYQPELPKC 261
 QY 193 KSPDVIYNSPIQG-KIYIKENRFQCKMGVEYERDDAVCTESEG-WRP-LPSC 244

Search completed: Thu Jun 8 21:43:01 2000
Job time : 30 secs.

 W39154

 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPerch_p protein - protein database search, using Smith-Waterman algorithm
 Run on: Thu Jun 8 21:44:51 2000; Maspar time 11.39 Seconds
 684.453 Million cell updates/sec
 Tabular output not generated.

Title: >US-09-316-163-11
 Description: (1-329) from US09316163.pep
 Perfect Score: 2539
 Sequence: 1 EDCNLPERRRNTTEILTGSMS.....PDIKHGGLYHEHMRPFVPV 329

Scoring table: PAM 150
 Gap 11
 188963 segs, 23686106 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: a-geneseq36
 I-geneseqp

Statistics: Mean 33.142; Variance 124.795; scale 0.266

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	1832	72.2	240	1	Human partial Compleme	3.40e-190
2	1550	65.0	216	1	Clone PRB9PF410 CFH r	1.47e-169
3	452	17.8	578	1	Human Complement facto	9.41e-33
4	433	17.1	581	1	Human C4 binding prote	1.05e-33
5	409	16.1	290	1	Human partial compleme	3.96e-31
6	409	16.1	778	1	Amino acid sequence of	3.96e-31
7	409	16.1	1930	1	Human complement recep	3.96e-31
8	406	16.1	2317	1	CR1 protein	8.29e-31
9	406	16.0	1537	1	Partial human compleme	8.29e-31
10	406	16.0	2039	1	Human complement type	1.56e-30
11	404	15.9	543	1	CR1-4 (85R, 87N) analo	1.74e-30
12	403	15.9	543	1	CR1-4 (99H, 103E) anal	2.22e-30
13	402	15.8	543	1	CR1-4 (114S) analogue.	3.64e-30
14	400	15.8	543	1	CR1-4 (64K) analogue.	4.65e-30
15	389	15.7	543	1	CR1-4 (78T, 79D) analo	4.65e-30
16	389	15.7	543	1	CR1-4 (369-376 SRKPPIC	4.65e-30
17	389	15.7	543	1	CR1-4 (52S, 53S, 54P)	4.65e-30
18	399	15.7	543	1	CR1-4 (364-367 NAAH) a	4.65e-30
19	399	15.7	543	1	CR1-4 (347T, 349Y) ana	5.95e-30
20	398	15.7	543	1	CR1-4 (109N, 110A, 111	5.95e-30
21	398	15.6	543	1	CR1-4 (92T) analogue.	7.61e-30
22	397	15.6	543	1	CR1-4 (121Q) analogue.	1.25e-29
23	395	15.6	543	1	CR1-4 (318-321 RNP) a	

ID	Score	Query Match	Length	DB ID	Description	Pred. No.
24	395	15.6	543	1	R28549	1.25e-29
25	395	15.6	543	1	R28563	1.25e-29
26	394	15.5	543	1	CR1-4 (117P) analogue.	1.59e-29
27	394	15.5	543	1	CR1-4 (115T) analogue.	1.59e-29
28	393	15.5	543	1	CR1-4 (65T) analogue.	2.04e-29
29	393	15.5	543	1	CR1-4 (37Y) analogue.	2.04e-29
30	393	15.5	543	1	CR1-4 (318R, 319N) ana	2.04e-29
31	393	15.5	543	1	CR1-4 (266-274 KLKQTN	2.04e-29
32	391	15.4	543	1	CR1-4 (116K) analogue.	3.33e-29
33	390	15.4	543	1	CR1-4 (35E) analogue.	4.26e-29
34	391	15.4	543	1	CR1-4 (57Y, 59K) analo	3.33e-29
35	389	15.3	543	1	CR1-4 (116K, 117P) ana	5.44e-29
36	387	15.2	543	1	CR1-4 (94H) analogue.	8.90e-29
37	386	15.2	543	1	CR1-4 (92T, 94H) analo	1.14e-28
38	385	15.2	543	1	CR1-4 (35E, 37Y) analo	1.45e-28
39	385	15.2	543	1	CR1-4 (114-117SRK, 12	3.04e-28
40	382	15.0	577	1	CR1-4 (44T, 47D, 49Y)	3.04e-28
41	382	15.0	577	1	Membrane co-factor pro	1.32e-27
42	376	14.8	830	1	P-selectin.	4.50e-27
43	371	14.6	263	1	Deduced sequence of co	4.06e-26
44	362	14.3	579	1	Amino acid sequence of	6.77e-24
45	341	13.4	302	1	Herpesvirus saimiri sc	

ALIGNMENTS

RESULT 1
 ID W39154 standard; Protein; 240 AA.
 AC W39154:
 DT 27-APR-1998 (first entry)
 DE Human partial Complement factor H protein fragment 1.
 KW Complement factor H; tumour associated antigen; renal cancer;
 KM urogenital cancer; medicament; modulator.
 OS Homo sapiens.
 PN W09738136-A1.
 PD 16-OCT-1997
 PE 09-APR-1997; US-05710
 PF 06-MAR-1997; US-812481.
 PR 09-APR-1996; US-015083.
 PR 09-APR-1996; US-630048.
 PR 06-MAR-1997; US-038614.
 PA (BARD-) BARD DIAGNOSTIC SCI INC.
 PI Enfield DL, Hass GM, Kinders RJ;
 DR WPT: 97-51742/47.
 DR N-PSDB; V02790.
 DT Treating or screening for cancer, e.g. renal or urogenital cancer -
 PT by modulating or detecting tumour associated human complement factor
 PS H related antigen, or nucleic acid encoding it
 CC Example 6B; Fig 6B; 104pp; English.
 CC This partial protein sequence represents a region of the human
 CC tumour-associated complement factor H (CFH). This sequence is used
 CC in the identification of complement factor H related proteins and
 CC antigens isolated from clone PRB9PF410 (see W39155). The detection of
 CC such proteins and a CFH antigens can be used in screening or for the
 CC treatment of renal or urogenital cancer, e.g. bladder, cervical or
 CC prostate cancer. Agents that may modulate this antigen could be used in
 CC the manufacture of a medicament for the treatment of a tumour cell.
 SQ Sequence 240 AA:
 Query Match 72.2%; Score 1832; DB 1; Length 240;
 Best Local Similarity 100.0%; Pred. No. 3.40e-190;
 Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 FTLTGNGVEFGYKAVYTCNCGYQLGELNRCBDTGTNDIPICEVYKCLPYAPENG 60
 |||
 QY 78 FTLTGNGVEFGYKAVYTCNCGYQLGELNRCBDTGTNDIPICEVYKCLPYAPENG 137
 |||
 Db 61 KIYSSAEPDREYHFGQAVFVNSGYKTEGDDEHMCSDGFMSKPKCVETISCKSPDY 120
 |||
 QY 138 KIYSSAEPDREYHFGQAVFVNSGYKTEGDDEHMCSDGFMSKPKCVETISCKSPDY 197
 |||
 Db 121 INGSPIQKIIYKENEFOYKCMNGYSEKGDVAVTESGWRPLPSCEKSCDNPIYPNG 180
 |||


```

OY 188 INGSPIKOKIYKENEREFOYKCNMGYSESGDAVCTESGMRPLPSCCEKSCDNPIYNG 257
Db 181 DYSPLRIKHTGDELTYYOORNGEYFATRGNTAKCTSTGIMAPRCLTKCDYDPIYHGL 240
OY 258 DYSPLRIKHTGDELTYYOORNGEYFATRGNTAKCTSTGIMAPRCLTKCDYDPIYHGL 317

RESULT 2
ID W39155 standard; Protein; 216 AA.
AC W39155;
DE 27-APR-1998 (first entry)
KW Complement factor H; tumour associated antigen; renal cancer;
  urogenital cancer; medicament; modulator.
OS Synthetic.
PS MO9738136-A1.
PE 16-OCT-1997.
PF 09-APR-1997; U05710
PR 06-MAR-1997; US-812481.
PR 09-APR-1996; US-015083.
PR 06-MAR-1997; US-038614.
PA (BARD-) BARD DIAGNOSTIC SCI INC.
PI Entfeld DL, Hass GM, Kinders RJ;
DR WPI: 97-512742/47.
DR N-PSDB: V02791.
PT Treating or screening for cancer, e.g. renal or urogenital cancer -
PT by modulating or detecting tumour associated human complement factor
PT H related antigen, or nucleic acid encoding it
PS Example 6B; Fig 6B; 104pp; English.
CC This partial protein is found in clone PR99FH410 and represents a
CC human tumour-associated complement factor H (CFH). The detection of this
CC protein and a CFH antigen can be used in screening or for the treatment
CC of renal or urogenital cancer, e.g. bladder, cervical or prostate cancer.
CC Agents that may modulate this antigen could be used in the manufacture of
CC a medicament for the treatment of a tumour cell.
SQ Sequence 216 AA;

Query Match 65.0%; Score 1650; DB 1; Length 216;
Best Local Similarity 99.5%; Pred. No. 1.47e-169;
Matches 215; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 TCNNGYLLGELINRREDTGGMTNDIPICEVYKCLPTAPBNGKIVSSAMPEDEYHGG 60
OY 95 TCNNGYLLGELINRREDTGGMTNDIPICEVYKCLPTAPBNGKIVSSAMPEDEYHGG 154
Db 61 AVRFVNSGKRIEGDEDMHSDGFMGKPKCYEISCKSPDVINGSPISOKIYKENER 120
OY 155 AVRFVNSGKRIEGDEDMHSDGFMGKPKCYEISCKSPDVINGSPISOKIYKENER 214
Db 121 FOYCKNMGYSESGDAVCTESGMRPLPSCCEKSCDNPIYNGDYSPLRIKHTGDEITY 180
OY 215 FOYCKNMGYSESGDAVCTESGMRPLPSCCEKSCDNPIYNGDYSPLRIKHTGDEITY 274
Db 181 QCRNGEYFATRGNTAKCTSTGIMAPRCLTKCDYDPIYHGL 216
OY 275 QCRNGEYFATRGNTAKCTSTGIMAPRCLTKCDYDPIYHGL 310

```

```

PR 06-OCT-1997; JP-272837.
PA (PROT-) PROTEGENE INC.
PA (SAGA) SAGAMI CHEM RES CENT.
PI Kato S, Sekine S;
DR WPI: 99-264019/22.
DR N-PSDB: X34737.
PT Human proteins with secretory signal sequences and nucleotide
PT sequences, useful in control of proliferation and differentiation of
PT cells
PS Claims 1; Page 55-58; 71pp; English.
CC This DNA encodes a protein having homology to human complement factor H,
CC which plays a role in the immunological mechanism involving the
CC complement reaction. The protein can also be used as an antigen for
CC preparing antibodies against the protein. The cDNA can be used as a probe
CC for gene diagnosis and the gene for gene therapy, as well as for large-
CC scale expression of the protein. The protein may also have immune
CC stimulating or suppressing activity, haematopoiesis regulating activity,
CC tissue growth activity, activity/inhibit activity, anti-inflammatory
CC activity, tumour inhibition activity, chemotactic/chemokinetic activity,
CC receptor/ligand activity, etc. The protein is identified by the presence
CC of a hydrophobic N-terminal secretory signal region, knowledge of the
CC cloning.
SQ Sequence 578 AA;

Query Match 17.8%; Score 452; DB 1; Length 578;
Best Local Similarity 32.8%; Pred. No. 9.41e-36;
Matches 82; Conservative 56; Mismatches 88; Indels 24; Gaps 20;

Db 54 YYCDNFVTPSGSYWDYIHCIDGMSPTVP-C-LRTCSKSDVEIENG-FIS---ESSSIY 107
OY 94 YTCNEGYQLL-GEI-NYRECDDGTNDIPICEVYKCL-PTAPBNGKIVSSAMPEDEYH 150
Db 108 ILNEETQYNKPGYATADGNSGSIITLQNG-WSTQ-PICIKF-CDMPVENSRAKSNGM 164
OY 151 HFGGAVRFVNSGKRIEGDEE--MHCSDGFMGKPKCYEISCKSPDVINGSPISOKI 207
Db 165 WFKLHDTIDYECYDSESSGNTTDSIVCGEDGWSHLPTCYNSSESGCPPPISNGDTTS 224
OY 208 IYKNERFQYKCNMGYSESGDAVCTESGMRPLPSC-EEK-SCDNPIYNGDYSPL 261
Db 225 FPOKVVYLPMSRVEYQCS-YTELQSKRYTCSNGSDSEPPRCISMKRCEPEIYHGLYX 283
OY 262 LRIRK-HRTGDEITYOORNGEYFATRGNTAKCTSTGIMAPRCLTKCDYDPIYHGLYH 319
Db 284 ENMRPPIYFPV 293
OY 320 ENMRPPIYFPV 329

RESULT 4
ID R13490 standard; Protein; 581 AA.
AC R13490;
DE 30-OCT-1991 (first entry)
KW Human C4 binding protein.
KW C4bp; monomer; complement protein; pI:OD.C4bp.3; SCR;
KW short consensus repeat.
OS Homo sapiens.
FH Key Location/Qualifiers
FT peptide 1..32
FT /label= signal_peptide 33..581
FT protein /label= C4bp 33..93
FT /label= SCR8 94..155
FT /label= SCR7 156..219
FT /label= SCR6 220..279
FT /label= SCR5 280..345
FT region /label= SCR4

```


FT region 346..406
 FT /label=SCR3
 FT region 407..464
 FT /label=SCR2
 FT region 465..523
 FT /label=SCR1
 FT domain 524..581
 FT /label=C4bp_core
 FT /note="responsible for multimer assembly"
 FT disulfide_bond 34..80
 FT /note="intradomain"
 FT disulfide_bond 65..92
 FT /note="intradomain"
 PN MO911461-A..
 PD 08-AUG-1991..
 PF 28-JAN-1991: U00567..
 PR 26-JAN-1990: US-470888..
 PA (BIOG-) BIOGEN INC..
 PI Paszek MP, Winkler G, Liu TR:
 DR WPI: 91-252613/34..
 DR N-PDB: Q13242..
 PT New C4 binding protein fusion proteins and DNA encoding them -
 PT comprise assemblies of C4bp monomers linked to functional moiety,
 PS e.g. A27, useful as delivery vehicles in diagnosis and therapy
 PS Example 1: Fig 1; 105pp; English.
 CC This sequence was deduced from human hepatocyte (Hep G2) cDNA
 CC obtained following PCR amplification. The protein is a monomer
 CC containing 8 SCRs. Each SCR forms a looped domain due to the
 CC presence of two intradomain disulphide bonds (only the disulphide
 CC bonds of SCR8 are labelled in the features table). Within each SCR,
 CC the first cysteine residue bonds with the third and the second
 CC cysteine residue bonds with the fourth. This secondary structure is
 CC responsible for the conformational flexibility of the C4bp monomer.
 CC The invention covers fusion proteins in which the monomer sequence,
 CC or a specified subfragment of it, e.g. having 5, 4, 3 or 1 SCR(s)
 CC is fused to the C-terminal of a protein such as a viral receptor,
 CC cell ligand, a bacterial, viral or parasitic immunogen, enzyme,
 CC cytokine, toxin, etc. See also Q13243-51.
 SQ Sequence 581 AA;

Query Match 17.1%; Score 433; DB 1; Length 581;
 Best Local Similarity 29.8%; Pred. No. 1.05e-33;
 Matches 97; Conservative 68; Mismatches 132; Indels 29; Gaps 26;

Db 51 TETFEKGTILKTYCLPGYRSHSTQTLTNSDGEWV-YNTFCILY-KR-CRHGELRLNGQ 107
 QY 20 SDQYTPPGTQAIYKRCRGY-RSLGNVIMGCR-GEWVALNPLKRCQKRRPGHGDPFGT 77
 Db 108 VEIKT-DL-SFGSQIEFSCEBEPFLIGSTSRCEVDRCGVGWSHPLPOCEIVKCPPTDI 165
 QY 78 FTLLGHWTFEYGVKAVYTCNGYOLLEINR-BC-DTD-GWTDNDIPICVYVACLPTAP 134
 Db 166 RNGR-HSG-E-ENFYAVGSEVTCSDPFRSLGSHASISCVENETIGVWRSPPTCEKI 221
 QY 135 ENGIIVSAMPPDEHYHGAQVAVFCNSGYKITEDEMKHS-D-D--GFSKKEPKVEI 190
 Db 222 TCRKPDVSHGEMVSGFGIYNYKDTIVFKCQKGFVLNGSSVINCADSKMNSPPACEPN 281
 QY 191 SKRSQDVINGSPISQ-KIITYKENERFOYKCMKGYEYSBRGAVC-TESGWRP-LPSCCEK 247
 Db 282 SCINLPDIPHASWETTPPTKEDYVVGTVLRKCHGPKYKTTDEPTTVCQKRLATPTPY 341
 QY 248 SCDN-PYIPNG-D-Y-S-PLRIK-HRTDELTTCOCNGEYPATRGNTA-KC-TSGWIPA 299
 Db 342 OGCEALCPEPKLNGEITQHRKSRP 367
 QY 300 PRCTIKPCDYIDIKRHGGLYHEMRMRP 325

RESULT 5
 ID W39156 standard; Protein; 290 AA.
 AC W39156;
 DT 27-APR-1998 (first entry)

DE Human partial complement factor H protein fragment 2.
 KW Complement factor H; tumour associated antigen; renal cancer;
 KW urogenital cancer; medicament; modulator.
 OS Homo sapiens.
 PN MO9738136-A1..
 PD 16-OCT-1997..
 PE 09-APR-1997: U05710..
 PR 06-MAR-1997: US-812481..
 PR 09-APR-1996: US-015083..
 PR 09-APR-1996: US-630048..
 PR 06-MAR-1997: US-038614..
 PA (BARD-) BARD DIAGNOSTIC SCI INC..
 PI Entfeld DL, Haas GM, Kinders RJ:
 DR WPI: 97-512742/47..
 DR N-PDB: V02792..
 PT Treating or screening for cancer, e.g. renal or urogenital cancer -
 PT by modulating or detecting tumour associated human complement factor
 PS H related antigen, or nucleic acid encoding it
 PS Example 6B, Fig 7B; 104pp; English
 CC This partial protein sequence represents a region of the human
 CC tumour-associated complement factor H (CFH). This sequence is used in the
 CC identification of complement factor H related proteins and antigens from
 CC clone PRBS3FH2910 (see W39157-W39159). The detection of this protein and
 CC a CFH antigen can be used in screening or for the treatment of renal or
 CC urogenital cancer, e.g. bladder, cervical or prostate cancer. Agents that
 CC may modulate this antigen could be used in the manufacture of a
 CC medicament for the treatment of a tumour cell.
 SQ Sequence 290 AA;

Query Match 16.1%; Score 409; DB 1; Length 290;
 Best Local Similarity 27.0%; Pred. No. 3.96e-31;
 Matches 75; Conservative 63; Mismatches 119; Indels 21; Gaps 20;

Db 1 GFRISENETTCYWGKW-S-SP-PQCEGLPCKSPPEISHGVAMMSQ-YOYGEVYTKC 56
 QY 37 GYRSLGNVIMCRKGEWVALNPLKRCQKRRPGHGDPFGFTLLGNGVEYGVKAVYTC 96
 Db 57 PEGFIDIDPALAK-CLGKMSHP-PSCIKTDCLSLSPENA-IPMGKK-D-VYKAGEQV 111
 QY 97 NEGYLLEINRECDTGTGWTNDIPICVVKCLPYTAPENCKIYSSAMEPDRHYHGOAV 156
 Db 112 TYTCATYVKMDGASNVTCINSR-WT-GRTCRDTCVNPPTVQNAVYRSMKYPSEGR 169
 QY 157 RYVCNSGKRYEEDDEMHOSDGFMSKKEPKVEIQC-KSPVINGSPIS-QKIITYKENER 214
 Db 170 VRYQCRSPYEMFGD-EYVCLNGNTEPPQCKDSTGCKGPPPLDNGDITFPLSYAPA 228
 QY 215 FOYKCNMGYE-YSERGDVAVCTESGWRPLPSCSEKS--CDNPPY-IPNGYSLRLIK-HRTG 269
 Db 229 SSVAYOCQN-LYQLBGNKRLTCRNQGWSEPPKC-LHPC 264
 QY 270 DETTYOCRNGEYPATRGNTAKCTSGWIPAPRCTLKPC 307

RESULT 6
 ID W73147 standard; protein; 778 AA.
 AC W73147;
 DT 29-JAN-1999 (first entry)
 DE Amino acid sequence of the soluble complement receptor 1 (SCR1).
 KW Human; soluble complement receptor 1; SCR1; T-cell; B-cell;
 KW mediated immune response; inhibition; tissue rejection; gene therapy;
 KW dystrophin; inflammatory response; interferon-gamma secretory response;
 KW autoimmune response; neurological response; Alzheimer's disease;
 KW Parkinson's disease; multiple sclerosis; systemic lupus erythematosus;
 KW rheumatoid arthritis; myasthenia gravis; epidermis bullousa;
 KW Hashimoto's disease.
 OS Homo sapiens.
 PN WO9845430-A1..
 PD 15-OCT-1998..
 PF 06-APR-1998: G01012..
 PR 05-APR-1997: GB-006950..
 PA (ANNE/) ANNEKOV A..
 PA (CHER/) CHERNAJOVSKY Y.

PI Annenkov A, Chernajovsky Y;
 DR WPI; 98-568350/48.
 PT Fragment of soluble human complement receptor 1 - useful for
 PT treating T-cell or B-cell mediated immune responses e.g.
 PT inflammatory responses such as rheumatoid arthritis
 PS Dislosure; Fig 1; 54pp; English.
 CC This is an amino acid sequence of the human soluble complement
 CC receptor 1 (SCR1), useful in the treatment of T-cell or B-cell
 CC mediated immune responses. It is used to inhibit a T-cell or
 CC B-cell-mediated immune response to prevent immune response-mediated
 CC tissue rejection and destruction or clearance or inactivation of an
 CC expressed protein especially from cells that have been treated by gene
 CC therapy to express the protein, e.g. dystrophin. The protein can also
 CC be used to inhibit a T-cell or B-cell-mediated inflammatory response,
 CC an interferon-gamma secretory response, autoimmune response or
 CC neurological response, e.g. Alzheimer's or Parkinson's disease or
 CC multiple sclerosis. Also the protein can be used to treat systemic
 CC lupus erythematosus, rheumatoid arthritis, myasthenia gravis,
 CC epidermis bulosa or Hashimoto's disease.
 SO Sequence 778 AA;

Query Match 16.1%; Score 409; DB 1; Length 778;
 Best Local Similarity 31.4%; Pred. No. 3,96e-31;
 Matches 97; Conservative 66; Mismatches 110; Indels 36; Gaps 30;

Db 68 PPIGTLYNECRPGYSGRPFST-ICLNKSVWTGAKD-R-CRRKSCRNPDPYNGMVAIK 124
 QY 24 YPEGTQAIYKRPYRISGNVIMWC-RKGEWVALNPKRCKRPGHGDPTFTTLTG 82
 Db 125 G-I-QFGSQIYKCTKGRILGSSSATCISGDTYIWDNETPICRIIPCGLPPTIT-NGD 181
 QY 83 GNVFEGYKAVYTCNEGYOLLGE--INRYEC-DIDGWTNDIPICEVAKC-LPVYAPENGK 138
 Db 182 FISTNRE-N-FHYGSVVTYRCNPGSGRKFELVGEPSIYCTSNDDQVIGMSGAPQCI 238
 QY 139 IVSSAMPEDEHYHGOAVRFVNCNGY--KI-E-GDEMHK-S-DD--GFWSKRPKCV 188
 Db 239 IPNCKTPPVNENGLIVSDNLSFLNEVVERQCPGFVMKPRRYKQALNKKEPELPSC 298
 QY 189 EIS-CKSPDYINGSPISQ-KIYKENRFQYKCMNGEYSEKGDVATE-SGWRP-LPSC 244
 Db 299 S-RVCOPEPVLAHERQKDNFSPQGEVYFSCPEG-Y-DLGAASMRCTPOGDWSPAA 355
 QY 245 EKSCKDNP-YIPNGDYSPL-RIKHRTGDEITTCRNGEYPATRGNTA-KCTSTG-WIPA- 299
 Db 356 PTCGVKSCD 364
 QY 300 PCTLKPCD 308

RESULT 7
 ID W45899 standard; peptide; 1930 AA.
 AC W45899;
 DT 30-JUN-1998 (first entry)
 DE Human complement receptor 1 (residues 1-1929).
 KW Membrane binding element; thrombotic disease; soluble protein;
 KW complement-related disease; integral membrane protein; inflammation;
 KM short consensus repeat; SCR 1-3; CRL; complement receptor type 1.
 OS Homo sapiens.
 FH Key
 FT Cross_links
 FT 1930
 FT /note= "Disulphide linked to Cys in peptide given
 in W45889"
 PN WO9802454-A2.
 PD 22-JAN-1998.
 PE 08-JUL-1997; E03715.
 PF 15-JUL-1996; GB-014871.
 PA (ADPR-) ADPROTECH PLC.
 PI Dodd I, Mossakowska DEI, Smith RAG;
 DR WPI; 98-110524/10.
 PT membrane binding groups - useful for treating complement-related
 PT thrombotic diseases, providing improved localisation at cellular

PT Membranes
 PS Claim 22; Pages 60-61; 75pp; English.
 CC This sequence represents human complement receptor 1 (CRL, CD 35)
 CC N-terminal fragment. The invention relates to a soluble derivative (A)
 CC of a soluble polypeptide (I), which comprises at least 2 heterologous
 CC membrane-binding elements (MBE) of low membrane affinity covalently
 CC associated with (I). MBE interact, independently and with thermodynamic
 CC additivity, with components of cellular or artificial membranes exposed
 CC to extracellular fluids. (A) are used to treat disorders treatable with
 CC (I) itself, specifically inflammation or any other complement-related
 CC disorder (e.g. neurological disease, graft rejection, myocardial
 CC infarction, sepsis, rheumatoid arthritis and many others; including
 CC application to indwelling devices) and thrombolytic disease, but also to
 CC treat allergy, induce weight loss, to treat ischemia or asthma and as
 CC immuno-modulators for treating multiple sclerosis. (A) are administered
 CC orally, topically, by injection or inhalation at 0.01-10 (preferably
 CC 0.1-10) mg/kg/day.
 SO Sequence 1930 AA;

Query Match 16.1%; Score 409; DB 1; Length 1930;
 Best Local Similarity 31.4%; Pred. No. 3,96e-31;
 Matches 97; Conservative 66; Mismatches 110; Indels 36; Gaps 30;

Db 22 PPIGTLYNECRPGYSGRPFST-ICLNKSVWTGAKD-R-CRRKSCRNPDPYNGMVAIK 78
 QY 24 YPEGTQAIYKRPYRISGNVIMWC-RKGEWVALNPKRCKRPGHGDPTFTTLTG 82
 Db 79 G-I-QFGSQIYKCTKGRILGSSSATCISGDTYIWDNETPICRIIPCGLPPTIT-NGD 135
 QY 83 GNVFEGYKAVYTCNEGYOLLGE--INRYEC-DIDGWTNDIPICEVAKC-LPVYAPENGK 138
 Db 136 FISTNRE-N-FHYGSVVTYRCNPGSGRKFELVGEPSIYCTSNDDQVIGMSGAPQCI 192
 QY 139 IVSSAMPEDEHYHGOAVRFVNCNGY--KI-E-GDEMHK-S-DD--GFWSKRPKCV 188
 Db 193 IPNCKTPPVNENGLIVSDNLSFLNEVVERQCPGFVMKPRRYKQALNKKEPELPSC 252
 QY 189 EIS-CKSPDYINGSPISQ-KIYKENRFQYKCMNGEYSEKGDVATE-SGWRP-LPSC 244
 Db 253 S-RVCOPEPVLAHERQKDNFSPQGEVYFSCPEG-Y-DLGAASMRCTPOGDWSPAA 309
 QY 245 EKSCKDNP-YIPNGDYSPL-RIKHRTGDEITTCRNGEYPATRGNTA-KCTSTG-WIPA- 299
 Db 310 PTCGVKSCD 318
 QY 300 PCTLKPCD 308

RESULT 8
 ID P92219 standard; protein; 2317 AA.
 AC P92219;
 DT 22-FEB-1990 (first entry)
 DE CRI protein
 KW Complement; cofactor.
 OS Homo sapiens (human).
 FH Key
 FT peptide
 FT 10.50
 FT /label= signal_peptide
 PN WO8909220-A.
 PD 05-OCT-1989.
 PE 31-MAR-1989; U01358.
 PR 01-APR-1988; US-176532.
 PA (TCEU) T Cell Sciences Inc; (WVJO) The Johns Hopkins University;
 PA (BRIG*) The Brigham and Women's Hospital.
 PI Fearon DT, Klickstein LB, Wong W, Carson G, Concino MF, Makrides SC;
 DR WPI; 89-309498/42.
 DR N-PSDB; N91477.
 PT New nucleic acid sequences encoding new CRI protein - and its fragment,
 PT for diagnosis and control of complement-related immune defects,
 PT inflammation, myocardial infarct, etc
 PS Claim 1; Fig. 1; 191pp; English.
 CC This is full-length CRI protein, and shortened forms are new, lacking
 CC the transmembrane region. The proteins and fragments bind C3b and/or

PI Human complement receptor type 1 gene

 WISEMAN (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

 Msearch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Thu Jun 8 21:45:24 2000; Msearch time 19.71 Seconds
 787.259 Million cell updates/sec
 Tabular output not generated.

Title: >US-09-316-163-11
 Description: (1-328) from US09316163.pep
 Perfect Score: 2539
 Sequence: 1 EDCEHLPPTRRNTEILTGSWS.....PDINHGGLYHNNRRPFV 329

Scoring table: PAM 150
 Gap 11

Searched: 142080 segs, 47172406 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: pir62
 1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 43.995; Variance 69.824; scale 0.630

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	2539	100.0	449	1 NBHUS	complement factor H P	0.00e+00
2	2539	100.0	1231	1 NBHUS	complement factor H P	0.00e+00
3	1772	69.8	1234	1 NBHUS	complement factor H P	0.00e+00
4	1336	48.7	669	2 S65551	factor H - bovine (fr	3.71e-284
5	603	23.7	1053	2 S46199	probable complement r	5.01e-80
6	447	17.6	597	2 S53711	C4BP alpha chain prec	8.45e-78
7	438	17.3	452	2 A35068	complement factor H-r	1.45e-76
8	433	17.1	303	2 H35068	apolipoprotein H-rela	4.26e-74
9	423	16.7	597	1 NBHUS	C4b-binding protein a	1.18e-70
10	423	16.7	2014	2 A28507	complement C3b/C4b re	1.18e-70
11	409	16.1	2039	2 A28507	complement C3b/C4b re	1.18e-70
12	409	16.1	2039	2 A28507	complement C3b/C4b re	1.18e-70
13	404	15.9	558	2 S57953	C4BP protein alpha ch	3.49e-69
14	404	15.9	661	2 I36937	complement receptor 1	6.13e-69
15	402	15.8	469	1 NBHUS	C4b-binding protein a	1.62e-65
16	388	15.3	560	2 T16833	hypothetical protein	1.62e-65
17	387	15.2	482	2 A34924	complement C3b/C4b re	2.84e-65
18	376	14.8	830	2 A30359	P-selectin precursor	7.21e-62
19	373	14.7	676	2 A45900	complement C3b recept	2.20e-61
20	371	14.6	345	1 NBHUS	apolipoprotein H prec	2.05e-60
21	367	14.5	768	2 I36821	P-selectin - rat	3.35e-59
22	362	14.3	579	2 A56740	sperm-egg recognition	

RESULT ENTRY	1	ALIGNMENTS
TITLE	NBHUS	#type complete
ALTERNATE_NAMES	complement factor H precursor, short splice form - human	
ORGANISM	complement factor H-related protein; complement protein H	
DATE	31-Dec-1993 #sequence_revision 23-Feb-1996 #text-change 22-Jun-1999	
ACCESSIONS	S03013; B60238; A27877; A61103; A26505; S10479	
REFERENCE	S00254	
#authors	Ripoche, J.; Day, A.J.; Harris, T.J.R.; Sim, R.B.	
#journal	Biochem. J. (1988) 249:593-602	
#title	The complete amino acid sequence of human complement factor H.	
#cross-references	MUID:88134059	
#accession	S03013	
#molecule-type	mRNA	
##residues	1-449 ##label RIP	
##cross-references	EMBL:X07523; EMBL:Y00716; NID:q32492; PDB:CAA30403.1; PDB:q758073	
#note	part of this sequence, including the amino end of the mature protein was confirmed by protein sequencing	
#note	402-Tyr was also found	
REFERENCE	A60238	
#authors	Estaller, C.; Schwaible, W.; Dierich, M.; Weiss, E.H.	
#journal	Eur. J. Immunol. (1991) 21:799-802	
#title	Human complement factor H: two factor H proteins are derived from alternatively spliced transcripts.	
#cross-references	MUID:91184292	
#accession	B60238	
##status	not compared with conceptual translation	
##molecule-type	mRNA	
##residues	1-33:434-449 ##label EST	
#note	only portions of this 1.8 kilobase mRNA were sequenced	
REFERENCE	A27877	
#authors	Schulz, T.F.; Schwaible, W.; Stanley, K.K.; Weiss, E.	
#journal	Dierich, M.P.	
#title	Human complement factor H: isolation of cDNA clones and partial cDNA sequence of the 38-kDa tryptic fragment containing the binding site for C3b.	
#cross-references	MUID:87054207	
#accession	A27877	
##molecule-type	mRNA	
##residues	1-55:401, 'Y', 403-449 ##label SCH	
#note	an additional nucleotide present within the codon for Glu-310 was thought to be a cloning artifact and was ignored in translation	


```

REFERENCE
#authors A61103
#journal Schwaible, W.; Zwirner, J.; Scholz, T.F.; Linke, R.P.;
#title Eur. J. Immunol. (1987) 17:1485-1489
#cross-references GDB:129095
#accession A61103
#status not compared with conceptual translation
#molecule_type mRNA
#residues 27-76 ##label sc2
#note this is a partial sequence of an alternatively spliced
1.8 kilobase mRNA that is translated to yield a 43 k
form related to factor H

REFERENCE
#authors A26505
#journal Sim, R.B.; Discipio, R.G.
#title Biochem. J. (1982) 205:285-293
#cross-references MUID:83048213
#accession A26505
#molecule_type protein
#residues 19-20,'Q',22-29,'V',31-33,'Q',35 ##label SIM
REFERENCE
#authors A44551
#journal Barlow, P.N.; Norman, D.G.; Steinkasserer, A.; Horne, T.J.;
#title Biochemistry (1992) 31:3626-3634
#cross-references MUID:92232649
#accession S10479
#molecule_type protein
#residues 19-20,'Q',22-29,'V',31-33,'Q',35 ##label SIM
#note annotation: NMR structure determination, residues 264-292
REFERENCE
#authors Kristensen, T.; Wetzel, R.A.; Tack, B.F.
#journal J. Immunol. (1986) 136:3407-3411
#title Structural analysis of human complement protein H: homology
with C4b binding protein, beta(2)-glycoprotein I, and the
Ba fragment of B.
#cross-references MUID:86169701
#accession S10479
#molecule_type mRNA
#residues 226-'M', 'Y', 403-449 ##label KRI
COMMENT ##cross-references GDB:M12383; NID:9180472; PIDN:AAA52013.1; PID:9180473
Factor H has also been found bound to cell membranes in an unknown
manner. However, it has at least one cell attachment site motif
in repeat 4.
COMMENT Alternative transcripts of 4.3, 1.8, and 1.4 kilobases are expressed
in liver. See also PIR:NBH0H.
GENETICS
#gene GDB:HFI, HF
#map_position 1932-1932 OMIM:134370
GENETICS
#gene GDB:HF2, HF
#map_position 1932-1932
#note the correspondence between the two loci and the sequences
indicated is unclear; factor H has been reported to have
several allelic forms
FUNCTION
#description a cofactor in the inactivation of C3b by serine proteinase I;
also increases the rate of dissociation of the C3bb
complex (C3 convertase) and the (C3b)nb complex (C5
convertase) in the alternative complement pathway
#superfamily complement factor H; complement factor H repeat
homology
#glycoprotein; plasma
KEYWORDS
#domain signal sequence #status predicted #label SIG\
#product complement factor H, short splice form #status
experimental #label MAT\
#domain complement factor H repeat homology #label FH01\
1-18
19-449
21-80

```

```

85-141
146-205
210-262
246-248
267-320
325-385
389-442
21-66,52-80,85-129,
114-141,146-192,
178-205,210-251,
237-262,267-309,
294-320,325-374,
357-385,389-431,
416-442
217
#disulfide_bonds #status predicted\
#binding_site carbohydrate (Asn) (covalent) #status
absent
SUMMARY
#length 449 #molecular_weight 51007 #checksum 6077
Query Match 100.0%; Score 2539; DB 1; Length 449;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 19 EDCNELPPRRNTEILGWSQDTPEGTQATYKCPGRSGNYIMVCRKGEWALNPLR 78
| 1 EDCNELPPRRNTEILGWSQDTPEGTQATYKCPGRSGNYIMVCRKGEWALNPLR 60
| 79 KQKRPCCGHPGDTPTGTTLFGNVFEYGVKAVYTCNKGYYLLGEINRECDTGTWNTDI 138
| 61 KQKRPCCGHPGDTPTGTTLFGNVFEYGVKAVYTCNKGYYLLGEINRECDTGTWNTDI 120
| 139 PICEVVKCLPTAPAPNGRIVSSAMEPDRHFHGOAVRVCSGKRIKGDDEMHCSDDGF 198
| 121 PICEVVKCLPTAPAPNGRIVSSAMEPDRHFHGOAVRVCSGKRIKGDDEMHCSDDGF 180
| 199 SEKKKCYEISCKSPDVINGSPISQKIIYKENERFOYKCNMGYSESRGDAVCTESGMR 258
| 181 SEKKKCYEISCKSPDVINGSPISQKIIYKENERFOYKCNMGYSESRGDAVCTESGMR 240
| 259 LPSCERKCDNPIYIPNGDYSPLRIKHRTGDEITYQCNGRFPATRGTAICTGTGWTAP 318
| 241 LPSCERKCDNPIYIPNGDYSPLRIKHRTGDEITYQCNGRFPATRGTAICTGTGWTAP 300
| 319 RCTLKPCDYPDIKKGGLYHNMRRRPFPV 347
| 301 RCTLKPCDYPDIKKGGLYHNMRRRPFPV 329
|
RESULT 2
ENTRY
TITLE NBH0H #type complete
ORGANISM complement factor H precursor, long splice form - human
DATE #formal_name Homo sapiens #common_name man
31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change
22-Jun-1999
ACCESSIONS S00254; A60238; A54726; A61565; A26505; I72654; S66298
REFERENCE
#authors Riipoche, J.; Day, A.J.; Harris, T.J.R.; Sim, R.B.
#journal Biochem. J. (1988) 249:593-602
#title The complete amino acid sequence of human complement factor
H.
#cross-references MUID:88134059
#accession S00254
#molecule_type mRNA
#residues 1-1231 ##label R1P
#note ##cross-references EMBL:Y00716; NID:931964; PIDN:CAA68704.1; PID:931965
402-Tyr was also found
#note parts of this sequence, including the amino and carboxyl
ends of the mature protein, were confirmed by protein
sequencing
REFERENCE
#authors A60238
#journal Estallier, C.; Schwaible, W.; Dierich, M.; Weiss, E.H.
#title Eur. J. Immunol. (1991) 21:799-802
Human complement factor H: two factor H proteins are derived
from alternatively spliced transcripts.

```


##cross-references MUID:91184292
 #accession A60238
 ##status not compared with conceptual translation
 ##molecule-type mRNA
 ##residues 1-56:1177-1231 #label EST
 #note only portions of this 4.3 kilobase mRNA were sequenced

REFERENCE
 #authors A54726
 Day, A.J.; Ripchoe, J.; Lyons, A.; McIntosh, B.; Harris, T.J.R.; Sim, R.B.
 #journal Biosci. Rep. (1987) 7:201-207
 #title Sequence analysis of a cDNA clone encoding the C-terminal end of human complement factor H.
 #cross-references MUID:88025472
 #accession A54726
 ##status not compared with conceptual translation
 ##molecule-type mRNA
 ##residues 'DFRN', 579-1231 #label DAY
 #cross-references GB:M17517; NID:9180497; PIDN:AAA52016.1; PID:9180498
 #note parts of this sequence were determined by protein sequencing

REFERENCE
 #authors A61565
 Ripchoe, J.; Day, A.J.; Willis, A.C.; Belt, K.T.; Campbell, R.D.; Sim, R.B.
 #journal Biosci. Rep. (1986) 6:65-72
 #title Partial characterization of human complement factor H by protein and cDNA sequencing: homology with other complement and non-complement proteins.
 #cross-references MUID:86188123
 #accession A61565
 ##status not compared with conceptual translation
 ##molecule-type mRNA
 ##residues 'METERMHLNAKI', 1050-1057, 'T', 1059-1102 #label R12

REFERENCE
 #authors A26505
 Sim, R.B.; Discolo, R.G.
 #journal Biochem. J. (1982) 205:285-293
 #title Purification and structural studies on the complement-system control protein beta-1-H (factor H).
 #cross-references MUID:83048213
 #accession A26505
 ##molecule-type protein
 ##residues 19-20, 'Q', 22-29, 'V', 31-33, 'Q', 35 #label SIM

REFERENCE
 #authors A4451
 Barlow, P.N.; Norman, D.G.; Steinkasserer, A.; Horne, T.J.; Pearce, J.; Driscoll, P.C.; Sim, R.B.; Campbell, I.D.
 #journal Biochemistry (1992) 31:3626-3634
 #title Solution structure of the fifth repeat of factor H: A second example of the complement control protein module.
 #cross-references MUID:92232649
 #contents annotation; NMR structure determination, residues 264-292
 #note A49224
 Norman, D.G.; Barlow, P.N.; Baron, M.; Day, A.J.; Sim, R.B.; Campbell, I.D.
 #journal J. Mol. Biol. (1991) 219:717-725
 #title Three-dimensional structure of a complement control protein module in solution.
 #cross-references MUID:91278097
 #contents annotation; NMR structure determination, residues 927-985
 #note I56100
 Esteller, C.; Koistinen, V.; Schwaeble, W.; Dietrich, M.P.; Weiss, E.H.
 #journal Immunol. (1991) 146:3190-3196
 #title Cloning of the 1.4-kb mRNA species of human complement factor H reveals a novel member of the short consensus repeat family related to the carboxy terminal of the classical 150-kDa molecule.
 #cross-references MUID:91201892
 #accession I72654
 ##status translated from GB/EMBL/DBJ
 ##molecule-type mRNA
 ##residues 1047-1231 #label RES
 #cross-references GB:M65294; NID:9183766; PIDN:AAA35948.1; PID:9183767
 #note S66298
 #authors Caron, J.A.; Bates, R.C.; Smith, A.I.; Tetroz, T.; Arellano,

#journal A.; Gordon, D.L.; Burns, G.F.
 #title Biochim. Biophys. Acta (1996) 1289:305-311
 #cross-references MUID:96205365
 #accession S66298
 ##status preliminary
 ##molecule-type protein
 ##residues 411-419:574-578, 580-582 #label CAR
 #note Factor H has also been found bound to cell membranes in an unknown manner. However, it has at least one cell attachment site motif in repeat 4.
 #comment Alternative transcripts of 4.3, 1.8, and 1.4 kilobases are expressed in liver. See also PIR:NBH0HS.
 #GENETICS
 #gene GDB:HF1; HF
 #cross-references GDB:120041; OMIM:134370
 #map-position 1q32-1q32
 #map-position 1q32-1q32
 #note the correspondence between the two loci and the sequences indicated is unclear; factor H has been reported to have several allelic forms

FUNCTION
 #description a cofactor in the inactivation of C3b by serine proteinase I; also increases the rate of dissociation of the C3bb complex (C3 convertase) and the (C3b)nb complex (C5 convertase) in the alternative complement pathway
 #cross-references MUID:129095
 #gene GDB:HF2; HF
 #map-position 1q32-1q32
 #note the correspondence between the two loci and the sequences indicated is unclear; factor H has been reported to have several allelic forms

CLASSIFICATION
 #pathway complement alternate pathway;
 #classification glycoprotein; plasma

KEYWORDS
 #domain signal sequence #status predicted #label SIG
 #product complement factor H #status experimental #label
 #product complement factor H, short splice form #status
 #product complement factor H repeat homology #label FH01
 #domain complement factor H repeat homology #label FH02
 #domain complement factor H repeat homology #label FH03
 #domain complement factor H repeat homology #label FH04
 #region cell attachment (R-G-D) motif
 #domain complement factor H repeat homology #label FH05
 #domain complement factor H repeat homology #label FH06
 #domain complement factor H repeat homology #label FH07
 #domain complement factor H repeat homology #label FH08
 #domain complement factor H repeat homology #label FH09
 #domain complement factor H repeat homology #label FH10
 #domain complement factor H repeat homology #label FH11
 #domain complement factor H repeat homology #label FH12
 #domain complement factor H repeat homology #label FH13
 #domain complement factor H repeat homology #label FH14
 #domain complement factor H repeat homology #label FH15
 #domain complement factor H repeat homology #label FH16
 #domain complement factor H repeat homology #label FH17
 #domain complement factor H repeat homology #label FH18
 #domain complement factor H repeat homology #label FH19
 #domain complement factor H repeat homology #label FH20

REFERENCE	author	journal	title
149728	Munoz-Canoves, P.; Tack, B.F.; Vlk, D.P.	Biochemistry (1989) 28:9891-9897	Analysis of complement factor H mRNA expression: Dexamethasone and IFN-gamma increase the level of H in L cells.
#cross-references	GB:M31979; NID:g193726; PIDN:AAA37762.1; PID:g193729		
#accession	149711		
#status	preliminary; translated from GB/EMBL/DBJ		
#molecule-type	mRNA		
#residues	1-18 ##label RES		
#cross-references	GB:M31979; NID:g193726; PIDN:AAA37762.1; PID:g193729		
#accession	149728		
#status	preliminary; translated from GB/EMBL/DBJ		
#molecule-type	mRNA		
#residues	1-19 ##label RE2		
#cross-references	GB:J02891; NID:g193805; PIDN:AAA37795.1; PID:g553926		
COMMENT	Two codominant alleles of factor H are present in mice. Factor H functions as a cofactor in the inactivation of C3b by serine proteinase I and also increases the rate of dissociation of the C3b complex (C3 convertase) and the (C3b)nb complex (C3 convertase) in the alternative complement pathway.		
GENETICS			
#map-position	1		
CLASSIFICATION	#superfamily complement factor H; complement factor H repeat homology		
KEYWORDS	complement alternate pathway; duplication; glycoprotein; plasma		
FEATURE			
1-18	#domain signal sequence #status predicted #label SIG\		
19-1234	#product complement factor H #status predicted #label MPT\		
21-80	#domain complement factor H repeat homology #label FH01\		
85-141	#domain complement factor H repeat homology #label FH02\		
146-205	#domain complement factor H repeat homology #label FH03\		
210-262	#domain complement factor H repeat homology #label FH04\		
246-248	#region cell attachment (R-G-D) motif\		
327-320	#domain complement factor H repeat homology #label FH05\		
325-385	#domain complement factor H repeat homology #label FH06\		
389-442	#domain complement factor H repeat homology #label FH07\		
448-505	#domain complement factor H repeat homology #label FH08\		
509-564	#domain complement factor H repeat homology #label FH09\		
569-622	#domain complement factor H repeat homology #label FH10\		
629-683	#domain complement factor H repeat homology #label FH11\		
690-743	#domain complement factor H repeat homology #label FH12\		
752-802	#domain complement factor H repeat homology #label FH13\		
808-861	#domain complement factor H repeat homology #label FH14\		
867-931	#domain complement factor H repeat homology #label FH15\		
936-989	#domain complement factor H repeat homology #label FH16\		
994-1048	#domain complement factor H repeat homology #label FH17\		
1053-1107	#domain complement factor H repeat homology #label FH18\		
1114-1168	#domain complement factor H repeat homology #label FH19\		
1172-1233	#domain complement factor H repeat homology #label FH20\		
21-66,52-80,85-129,			
114-141,146-192,			
178-205,210-251,			
237-262,267-309,			
294-320,325-374,			
357-385,389-431,			
416-442,448-494,			
477-505,509-553,			
536-564,569-610,			
597-622,629-672,			
658-683,690-732,			
718-743,752-791,			
780-802,808-850,			
836-861,867-920,			
906-931,936-978,			
964-989,994-1037,			
1023-1048,			
1053-1096,			
1082-1107,			


```

Db 219 GNVPAIRKENDVLRHCDRAFKHIDRSPYCIKQIKAEWSPPLCESIKRLTIMDG 278
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 200 GSPISQIKTIYKENERFOYKCNMGEYSEKRGDA-V--CTESSGRPLPSCSEKSCDNPYIPN 256
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 279 TRREAVANLPSPEGLTAVICARTSMISIPQTSVTTTCQDNGEWSIRPTCEVRC 334
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 257 GDYSL-RIKHRTGDEITYQC-RNGFYPATRGNT--AKCTSTG-WIPAPRTCLKPC 307
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
ENTRY 553711 #type complete
TITLE C4BP alpha chain precursor - rabb1t
ORGANISM #formal_name Oryctolagus cuniculus #common_name domestic
          #abbt
DATE 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change
      10-Sep-1999
ACCESSIONS 553711
REFERENCE 553711
#authors de Frutos, P.G.; Dahlbaeck, B.
#journal Biochim. Biophys. Acta (1995) 1261:285-289
#title CDNA structure of rabbit C4b-binding protein alpha-chain.
        Preserved sequence motif in complement regulatory protein
        modules which bind C4b.
#cross-references MUID:95226458
#accession S53711
#status preliminary; nucleic acid sequence not shown
#molecule_type mRNA
#residues 1-597 #label DEF
#cross-references EMBL:Z45490
CLASSIFICATION #superfamily C4b-binding protein alpha chain; complement
                factor H repeat homology

FEATURE
50-107 #domain complement factor H repeat homology #label FH1\
112-169 #domain complement factor H repeat homology #label FH2\
174-234 #domain complement factor H repeat homology #label FH3\
239-294 #domain complement factor H repeat homology #label FH4\
364-422 #domain complement factor H repeat homology #label FH5\
426-480 #domain complement factor H repeat homology #label FH6\
484-538 #domain complement factor H repeat homology #label FH7\
SUMMARY #length 597 #molecular_weight 66130 #checksum 6473

Query Match 17.6%; Score 447; DB 1; Length 597;
Best Local Similarity 30.9%; Pred. No. 5,01e-80;
Matches 101; Conservative 62; Mismatches 133; Indels 31; Gaps 26;

Db 66 SENEYGTGTLTKTCRPGTIRNGLNPLITCKPGLW-SYDTF--CVKKRCRNGDLPNGQ 122
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 20 SDQTYGEGTQAIKCRPRGRNSLG-NVIMWCR-KGEWVALNPLKCKQKRPCHGDPFGT 77
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 VEYKT-D-FSFGSQIEFSCSEGYILIGSTT-SHCDDIQEGKGVESDPLKCEIYKCEPPN 179
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 78 FTLTGQNVFEYGVKAYVITCNESQQLGELIYRCDT-D-G--WTNDIPICEVVKCLPYTA 133
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 180 IINGK-HNGGN-DIH-TYGSSTVYSCNPRFSLDGEASISCTYKNTVGVSPSPVCKE 236
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 134 PENGKIVSSAMEPDRHYHGOAVRFVNCNGYKTEGDEMHC--DD--GWSKSKKCYE 189
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 237 IICSPNVPHGKILISGFPIYNNKDSIMTCTIDGFLVLRSSSLHCLDLSKWNPSPVCS 296
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 190 ISCKSPDVINGSIISO-KITLYENEFQYKCNMGEYSEKRGDAVCT-ESGWRLPS-CEE 246
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 297 NSCLGLPNVPHASQOQYQWSTKEGVSVGTETLRKCRPGYRPVADPPIVTCQEDLRWSP 356
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 247 KSC-DNPIYIPNG-D-Y-S-PLRIK-HRTGDEITYQC-RNGFYPATRGNT-AKCTST-GWIP 298
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 357 FAGCEAICCPHQDNGAL-TEHSRNP 382
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 299 APRCTLKPCDYPDIKHGGLYHENMRP 325
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
ENTRY A35068 #type fragment

```

```

TITLE complement factor H-related protein 3A4/5G4 - mouse
      (fragment)
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 27-Jul-1990 #sequence_revision 31-Dec-1993 #text_change
      20-Aug-1999
ACCESSIONS A35068; B35068; C35068; D35068; E35068; F35068; G35068
REFERENCE A35070
#authors ViK, D.P.; Munoz-Canoves, P.; Kozono, H.; Martin, L.G.; Tack,
        B.F.; Chaplin, D.D.
#journal J. Biol. Chem. (1990) 265:3193-3201
#title Identification and sequence analysis of four complement
        factor H-related transcripts in mouse liver.
#cross-references MUID:90153965
#accession A35068
#molecule_type mRNA
#residues 1-452 #label VIK
#cross-references GB:M29010; NID:g192561; PIDN:AAA37415.1; PID:g387128
#note translation of the nucleotide sequence is not complete
CLASSIFICATION #superfamily complement factor H; complement factor H repeat
                homology

FEATURE
28-61 #domain complement factor H repeat homology #label FH1\
86-146 #domain complement factor H repeat homology #label FH2\
150-203 #domain complement factor H repeat homology #label FH3\
209-266 #domain complement factor H repeat homology #label FH4\
270-325 #domain complement factor H repeat homology #label FH5\
332-386 #domain complement factor H repeat homology #label FH6\
390-451 #domain complement factor H repeat homology #label FH7\
SUMMARY #length 452 #checksum 2383

Query Match 17.3%; Score 438; DB 2; Length 452;
Best Local Similarity 56.4%; Pred. No. 8.46e-78;
Matches 53; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

Db 15 TALLSTAKGEKTCSPYIINGITPTRHRIHSDDIRRCNGFYPTGVSTKCTPTG 74
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 236 SGWRPLPSCSEKSCDNPYIPNGDYSPLRIKHRTGDEITYCNRNGFYPATRGNTAKCTSTG 295
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 75 WIPVPRCTLKPCDFEPOFKYGRLYEESLRPNPV 108
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 296 WIPAPRTCLKPCDYPDIKHGGLYHENMRPYPV 329
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
ENTRY H35068 #type complete
TITLE apolipoprotein H-related protein 23L1 - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 27-Jul-1990 #sequence_revision 31-Dec-1993 #text_change
      10-Sep-1997
ACCESSIONS H35068; A35069; B35069; I35068
REFERENCE A35070
#authors ViK, D.P.; Munoz-Canoves, P.; Kozono, H.; Martin, L.G.; Tack,
        B.F.; Chaplin, D.D.
#journal J. Biol. Chem. (1990) 265:3193-3201
#title Identification and sequence analysis of four complement
        factor H-related transcripts in mouse liver.
#cross-references MUID:90153969
#accession H35068
#status preliminary
#molecule_type mRNA
#residues 1-303 #label VIK
#cross-references GB:M29007; NID:g192557; PID:g309164; GB:J05259
#note translation of the nucleotide sequence is not complete
CLASSIFICATION #superfamily complement factor H repeat homology
                duplication
KEYWORDS
FEATURE
28-61 #domain complement factor H repeat homology #label FH01\
86-146 #domain complement factor H repeat homology #label FH02\
150-203 #domain complement factor H repeat homology #label FH03\
212-266 #domain complement factor H repeat homology #label FH04\
SUMMARY #length 303 #molecular_weight 34498 #checksum 2835

Query Match 17.1%; Score 433; DB 2; Length 303;

```


Fri Jun 9 10:53:40 2000

US-09-316-163-11.rpr

Page 12

QY 293 STGWI-PAPROCTLKPCDYPDIKHGGLYHEN 321

Search completed: Thu Jun 8 21:45:47 2000
Job time : 23 secs.

 M P E R E H
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

Mpsrch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Thu Jun 8 21:46:04 2000; Maspar time 12.45 Seconds
 804,485 Million cell updates/sec
 Tabular output not generated.

Title: >US-09-316-163-11
 Description: (1-329) from US09316163.pep
 Perfect Score: 2539
 Sequence: 1 EDCNEUPPRNTEILGSMS.....PIKHGGLYHNMRRPYFPV 329

Scoring table: PAM 150
 Gap 11

Searched: 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: swiss-prot38
 1:swissprot

Statistics: Mean 44.764; Variance 62.937; scale 0.711

Pred. No. 18, the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2539	100.0	1231	1	CEFAH_HUMAN COMPLEMENT FACTOR H PR	0.00e+00
2	1772	69.8	1234	1	CEFAH_MOUSE COMPLEMENT FACTOR H PR	0.00e+00
3	433	17.1	1597	1	C4BP_HUMAN C4B-BINDING PROTEIN AL	1.94e-86
4	409	16.1	2039	1	CRL_HUMAN COMPLEMENT RECEPTOR TY	8.70e-80
5	404	15.9	538	1	C4BP_RAT C4B-BINDING PROTEIN AL	2.09e-78
6	388	15.8	469	1	C4BP_MOUSE C4B-BINDING PROTEIN PR	7.44e-78
7	382	15.3	263	1	VCP_VACCV COMPLEMENT CONTROL PRO	5.28e-74
8	376	14.8	830	1	LEM3_HUMAN P-SELECTIN PRECURSOR (1.02e-70
9	371	14.6	345	1	APOH_MOUSE BETA-2-GLYCOPROTEIN I	2.38e-69
10	367	14.5	768	1	LEM3_RAT P-SELECTIN PRECURSOR (2.94e-68
11	363	14.3	668	1	F13B_MOUSE COAGULATION FACTOR XII	3.61e-67
12	357	14.1	610	1	C4BP_BOVIN C4B-BINDING PROTEIN AL	1.54e-65
13	343	13.5	768	1	LEM3_MOUSE P-SELECTIN PRECURSOR (9.54e-62
14	341	13.4	360	1	CCPH_HYVSA COMPLEMENT CONTROL PRO	3.31e-61
15	340	13.4	661	1	F13B_HUMAN COAGULATION FACTOR XII	6.15e-61
16	334	13.2	345	1	APOH_BOVIN BETA-2-GLYCOPROTEIN I	2.54e-59
17	333	13.2	377	1	MCP_HUMAN MEMBRANE COPROTEIN I	4.72e-59
18	333	13.1	345	1	APOH_HUMAN BETA-2-GLYCOPROTEIN I	1.44e-55
19	320	12.6	620	1	LEM2_HUMAN E-SELECTIN PRECURSOR (5.75e-54
20	314	12.4	612	1	LEM2_MOUSE E-SELECTIN PRECURSOR (6.68e-53
21	310	12.2	958	1	HIG_DROME LOCOMOTION-RELATED PRO	7.71e-52
22	306	12.1	345	1	APOH_CANFA BETA-2-GLYCOPROTEIN I	2.27e-52
23	308	12.1	484	1	LEM2_PIG E-SELECTIN PRECURSOR (

24	306	12.1	551	1	LEM2_RABIT E-SELECTIN PRECURSOR (7.71e-52
25	306	12.1	769	1	LEM3_SHEEP P-SELECTIN PRECURSOR (7.71e-52
26	302	11.9	1033	1	CR2_HUMAN COMPLEMENT RECEPTOR TY	8.86e-51
27	299	11.8	330	1	CEFAH_HUMAN COMPLEMENT FACTOR H-LI	5.51e-50
28	299	11.8	549	1	LEM2_RAT E-SELECTIN PRECURSOR (5.51e-50
29	299	11.8	611	1	LEM2_CANFA E-SELECTIN PRECURSOR (5.51e-50
30	293	11.5	381	1	DAF_HUMAN COMPLEMENT DECAV-ACCEL	2.11e-48
31	289	11.4	1025	1	CR2_MOUSE COMPLEMENT RECEPTOR TY	2.38e-47
32	284	11.2	340	1	DAF_PONPY COMPLEMENT DECAV-ACCEL	4.89e-46
33	281	11.1	331	1	CFHD_HUMAN COMPLEMENT FACTOR H-LI	2.98e-45
34	280	11.0	507	1	DAF_CAVPO COMPLEMENT DECAV-ACCEL	5.44e-45
35	276	10.9	297	1	APOH_RAT BETA-2-GLYCOPROTEIN I	6.03e-44
36	276	10.9	485	1	LEM2_BOVIN E-SELECTIN PRECURSOR (6.03e-44
37	276	10.9	1019	1	LFC_TACTR LIMBUS CLOTTING FACTO	6.03e-44
38	271	10.7	646	1	LEM3_BOVIN P-SELECTIN PRECURSOR (1.91e-42
39	254	10.0	390	1	DAF1_MOUSE COMPLEMENT DECAV-ACCEL	2.95e-38
40	238	9.4	407	1	DAF2_MOUSE COMPLEMENT DECAV-ACCEL	3.52e-34
41	232	9.1	372	1	LEM1_RAT L-SELECTIN PRECURSOR (1.15e-32
42	231	9.1	372	1	LEM1_MOUSE L-SELECTIN PRECURSOR (1.15e-32
43	220	8.7	372	1	LEM1_PANTR L-SELECTIN PRECURSOR (1.15e-29
44	220	8.7	372	1	LEM1_MACMU L-SELECTIN PRECURSOR (1.15e-29
45	220	8.7	372	1	LEM1_HUMAN L-SELECTIN PRECURSOR (1.15e-29

ALIGNMENTS

RESULT ID	1	CEFAH_HUMAN	STANDARD:	PRT:	1231 AA.
AC	P08603:				
DT	01-AUG-1988 (Rel. 08, Created)				
DT	01-JUN-1990 (Rel. 13, Last sequence update)				
DT	15-FEB-2000 (Rel. 39, Last annotation update)				
DE	COMPLEMENT FACTOR H PRECURSOR.				
GN	HPI OR HP OR CFH.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
RN	[1]	SEQUENCE FROM N.A.			
RC	TISSUE=LIVER;				
RX	MEDLINE: 88134059.				
RA	Rapoche J., Day A.J., Harris T.J.R., Sim R.B.;				
RT	"The complete amino acid sequence of human complement factor H.;"				
RL	Biochem. J. 249:593-602(1988).				
RN	[2]	SEQUENCE OF 53-445 FROM N.A.			
RP	MEDLINE: 87054207.				
RA	Schulz T.F., Schwaeble W., Stanley K.K., Weiss E., Dierich M.P.;				
RT	"Human complement factor H: Isolation of cDNA clones and partial cDNA sequence of the 38-kDa tryptic fragment containing the binding site for C3b.;"				
RL	Eur. J. Immunol. 16:1351-1355(1986).				
RN	[3]	SEQUENCE OF 226-445 FROM N.A., AND PARTIAL SEQUENCE.			
RP	MEDLINE: 86169701.				
RA	Kristensen T., Wetzel R.A., Jack B.F.;				
RT	"Structural analysis of human complement protein H: homology with C4b binding protein, beta 2-glycoprotein I, and the Ba fragment of B2.;"				
RL	J. Immunol. 136:3407-3411(1986).				
RN	[4]	SEQUENCE OF 1047-1231 FROM N.A.			
RP	MEDLINE: 91201892.				
RA	Estallier C., Koistinen V., Schwaeble W., Dierich M.P., Weiss E.H.;				
RT	"Cloning of the 1.4-kb mRNA species of human complement factor H reveals a novel member of the short consensus repeat family related to the carboxy terminal of the classical 150-kDa molecule.;"				
RL	J. Immunol. 146:3190-3196(1991).				
RN	[5]	SEQUENCE OF 19-35.			
RP	MEDLINE: 83048213.				
RA	Sim R.B., Dischlo R.G.;				
RT	"Purification and structural studies on the complement-system control protein beta 1H (Factor H).;"				

[illegible]

```

OY      300  PROTLKPCDYPDIKGLGYHENMRAP 325

RESULT      4
ID          CRI_HUMAN      STANDARD;      PRT;      2039 AA.
AC          P19727;
DT          01-NOV-1990 (Rel. 16, Created)
DT          01-NOV-1990 (Rel. 16, Last sequence update)
DT          01-NOV-1997 (Rel. 35, Last annotation update)
DE          COMPLEMENT RECEPTOR TYPE 1 PRECURSOR (C3B/C4B RECEPTOR) (CD35
DE          ANTIGEN).
OS          CRI OR C3BR.
OS          Homo sapiens (Human).
OC          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC          Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN          [1]
RP          SEQUENCE FROM N.A.
RP          MEDLINE; 89035992.
RA          Klickstein L.B., Bartow T.J., Miletic V., Rabson L.D., Smith J.A.,
RA          Fearon D.T.;
RL          "Identification of distinct C3b and C4b recognition sites in the
RT          human C3b/C4b receptor (CRI, CD35) by deletion mutagenesis.";
RL          J. Exp. Med. 168:1699-1717(1988).
RN          [2]
RN          SEQUENCE OF 503-2039 FROM N.A.
RX          MEDLINE; 87168191.
RA          Klickstein L.B., Wong W.W., Smith J.A., Weis J.H., Wilson J.G.,
RA          Fearon D.T.;
RT          "Human C3b/C4b receptor (CRI). Demonstration of long homologous
RT          repeating domains that are composed of the short consensus repeats
RT          characteristics of C3/C4 binding proteins.";
RT          J. Exp. Med. 165:1095-1112(1987).
RN          [3]
RN          SEQUENCE OF 761-783; 831-845 AND 1179-1195 FROM N.A.
RP          MEDLINE; 86067975.
RA          Wong W.W., Klickstein L.B., Smith J.A., Weis J.H., Fearon D.T.;
RT          "Identification of a partial cDNA clone for the human receptor for
RT          complement fragments C3b/C4b.";
RT          Proc. Natl. Acad. Sci. U.S.A. 82:7711-7715(1985).
CC          -I- FUNCTION: CRI, PRESENT ON ERYTHROCYTES, LEUCOCYTES, GLOMERULAR
CC          PODOCYTES, AND SPLENIC FOLLICULAR DENDRITIC CELLS, MEDIATES THE
CC          BINDING BY THESE CELLS OF PARTICLES AND IMMUNE COMPLEXES THAT HAVE
CC          ACTIVATED COMPLEMENT.
CC          .
CC          -I- SUBUNIT: MONOMER.
CC          -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC          -I- POLYMORPHISM: CRI IS RESPONSIBLE FOR THE KNOBS BLOOD GROUP SYSTEM.
CC          -I- MISCELLANEOUS: SEVEN SHORT CONSENSUS REPEATS (SCR) CONSTITUTE A
CC          LONG HOMOLOGOUS REPEAT (LHR). THE N-TERMINAL TWO SRCS OF LHR-A
CC          CONTAINED A SITE DETERMINING CA SPECIFICITY, AND THE N-TERMINAL
CC          TWO SRCS OF LHR-B AND -C EACH HAD A SITE DETERMINING C3
CC          SPECIFICITY.
CC          -I- MISCELLANEOUS: THIS IS THE SEQUENCE OF THE F ALLOTYPE OF CRI.
CC          -I- SIMILARITY: CONTAINS 30 SUSHI (SCR) REPEATS.
CC          -----
CC          THIS SWISS-PROT entry is copyright. It is produced through a collaboration on
CC          between the Swiss Institute of Bioinformatics and the EMBL outstation
CC          the European Bioinformatics Institute. There are no restrictions on
CC          use by non-profit institutions as long as its content is in no way
CC          modified and this statement is not removed. Usage by and for commercial
CC          entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC          or send an email to license@isb-sib.ch).
CC          -----
DR          EMBL; M1569; AAA52297.1; -
DR          EMBL; M1617; AAA52298.1; -
DR          EMBL; M1618; AAA52299.1; -
DR          EMBL; Y00816; CAA68755.1; -
DR          EMBL; X05309; CAA28933.1; -
DR          PIR; A28507; A28507.
DR          PIR; A24748; A24748.
DR          PIR; B24748; B24748.
DR          PIR; C24748; C24748.
DR          PIR; S03843; S03843.

```


RL address and inflammation.";
 RN Cell 56:1033-1044(1989).
 RP [2]
 RP STRUCTURE BY NMR OF 160-199.
 RX MEDLINE: 97057176.
 RA Freedman S.J., Sanford D.G., Bachovchin W.W., Furie B.C., Baleja J.D.,
 RT Furie B.;
 RT "Structure and function of the epidermal growth factor domain of P-
 RT selectin.";
 RL Biochemistry 35:13733-13744(1996).
 RP [3]
 RP 3D-STRUCTURE MODELING OF 42-161.
 RX MEDLINE: 94093388.
 RA Bajorath J., Stenkamp R., Aruffo A.;
 RT "Knowledge-based model building of proteins: concepts and examples.";
 RL Protein Sci. 2:1798-1810(1993).
 RN [4]
 RP VARIANTS ASN-331: ASP-603: VAL-640 AND PRO-756.
 RX MEDLINE: 98334547.
 RA Heitmann S.M., Ricard S., Nicoud V., Mallet C., Evans A.,
 RA Rudavals J.B., Arveiler D., Luc G., Cambien F.;
 RT "The P-selectin gene is highly polymorphic; reduced frequency of the
 RT Pro715 allele carriers in patients with myocardial infarction.";
 RL Hum. Mol. Genet. 7:1277-1284(1998).
 CC -I- FUNCTION: CA12(+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS
 CC TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
 CC INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
 CC LEUKOCYTES. THE LIGAND RECOGNIZED IS STALL1-LEMIS X.
 CC -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -I- TISSUE SPECIFICITY: STORED IN THE ALPHA-GRANULES OF PLATELETS
 CC AND WEIBEL-PALADE BODIES OF ENDOTHELIAL CELLS. UPON CELL
 CC ACTIVATION BY AGONISTS, P-SELECTIN IS TRANSPORTED RAPIDLY TO
 CC THE CELL SURFACE.
 CC -I- SIMILARITY: TO OTHER SELECTINS/LECAMS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -I- SIMILARITY: CONTAINS 9 SUSHI (SCR) REPEATS.
 CC -I- DATABASE: NAME=PROW; NOTE=CD guide CD62p entry;
 CC WWW:"http://www.ncbi.nlm.nih.gov/prow/cd/cd62p.htm".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC -----
 DR EMBL: M60234; AAA35910.1; JOINED.
 DR EMBL: M60217; AAA35910.1; JOINED.
 DR EMBL: M60218; AAA35910.1; JOINED.
 DR EMBL: M60219; AAA35910.1; JOINED.
 DR EMBL: M60222; AAA35910.1; JOINED.
 DR EMBL: M60223; AAA35910.1; JOINED.
 DR EMBL: M60224; AAA35910.1; JOINED.
 DR EMBL: M60225; AAA35910.1; JOINED.
 DR EMBL: M60226; AAA35910.1; JOINED.
 DR EMBL: M60227; AAA35910.1; JOINED.
 DR EMBL: M60228; AAA35910.1; JOINED.
 DR EMBL: M60229; AAA35910.1; JOINED.
 DR EMBL: M60231; AAA35910.1; JOINED.
 DR EMBL: M60232; AAA35910.1; JOINED.
 DR EMBL: M60233; AAA35910.1; JOINED.
 DR EMBL: M25322; -; NOT_ANNOTATED_CDS.
 DR PIR: A30359; A30359.
 DR PDB: 1FSB; 01-APR-97.
 DR PDB: 1KID; 03-APR-96.
 DR MIM: 173610; -
 DR PRINTS: PS00343; SELECTIN.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; 1.
 DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.

DR PFAM; PF00008; EGF; 1.
 DR PFAM; PF00059; Lectin; C; 1.
 DR PFAM; PF00084; sushi; 9.
 KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
 KW Selectin; signal; Sushi; Repeat; Polymorphism; 3d-structure.
 FT SIGNAL 1 41
 FT CHAIN 42 830
 FT DOMAIN 42 771
 FT TRANSMEM 772 795
 FT DOMAIN 796 830
 FT DOMAIN 58 158
 FT DOMAIN 159 195
 FT DOMAIN 199 762
 FT REPEAT 199 258
 FT REPEAT 261 320
 FT REPEAT 323 382
 FT REPEAT 385 444
 FT REPEAT 447 506
 FT REPEAT 509 568
 FT REPEAT 571 630
 FT REPEAT 641 700
 FT REPEAT 703 762
 FT DISULFID 60 158
 FT DISULFID 131 150
 FT DISULFID 163 174
 FT DISULFID 168 183
 FT DISULFID 185 194
 FT DISULFID 200 244
 FT DISULFID 230 257
 FT DISULFID 262 306
 FT DISULFID 292 319
 FT DISULFID 324 368
 FT DISULFID 354 381
 FT DISULFID 386 430
 FT DISULFID 416 443
 FT DISULFID 448 478
 FT DISULFID 510 554
 FT DISULFID 540 567
 FT DISULFID 572 616
 FT DISULFID 602 629
 FT DISULFID 642 686
 FT DISULFID 672 699
 FT DISULFID 704 748
 FT DISULFID 734 761
 FT CARBOHYD 54 54
 FT CARBOHYD 98 98
 FT CARBOHYD 180 180
 FT CARBOHYD 212 212
 FT CARBOHYD 219 219
 FT CARBOHYD 411 411
 FT CARBOHYD 460 460
 FT CARBOHYD 518 518
 FT CARBOHYD 665 665
 FT CARBOHYD 716 716
 FT CARBOHYD 723 723
 FT CARBOHYD 741 741
 FT SITE 818 821
 FT VARIANT 331 331
 FT VARIANT 603 603
 FT VARIANT 640 640
 FT VARIANT 756 756
 FT SEQUENCE 830 AA; 90844 MM; 90844 MM; 90844 MM;
 Query Match 14.8%; Score 376; DB 1; Length 830;
 Best Local Similarity 28.8%; Pred. No. 1 02e-70;
 Matches 85; Conservative 52; Mismatches 135; Indels 23; Gaps 21;

Db 343 AFAYGSSCKEPCQPGYRVLGMDIRCTIDSGHWSA--PLPTCEALSC-EPLESPVHGSMDC 399
 QY 23 TYPGTAQYKRCRPGYSLGNVWVCRK-GEWVALNLFRCKQKRPCHPDPTF-GTFLL 80
 Db 400 SPSLRAFYQYDNCSEFRCAGFGLRGADIVR-CDNLGQWTPAPVCAALQCOQDLPVFNAR 458
 QY 81 TGG-NVEYGVKAVYTNNEGVLGELTYRECDTDG-WTNDIPICEVVKLPYAPENGR 138
 Db 459 -VNCS-HPGFAFRY-QSVCFTCNESGLLVGASVLCATGNMNSVPECOAIPCPFLS 515
 QY 139 IVSSAMPDEREYHGQAV-RFVCSGYKIEGDEMHCSDDGFWKSKRPKVEISCKSP-D 196
 Db 516 PONGTMCVCPPLGSSSKYSKSCFCPCDBGYSLSGPERLDCITRSGRWDSPPMCAIKCPFL 575
 QY 197 VINSPLS-OKITKENERF-QYKCMNGEYSERGAVCESG-WRPLPS-CEKSCDNP 252
 Db 576 FAPGSGSLDSDTRGEFNVGSTCFSCNNGFKLEGPNNV-CTTSGWMSATPTTC 629
 QY 253 YIPN-G-DYSPLRKIRHTGDELTTCRNGFYPATKNTAKTSTG-WIPAPR-C 302
 RESULT 9
 ID APOH_MOUSE STANDARD; PRT; 345 AA.
 AC 001339;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE BETA-2-GLYCOPROTEIN I PRECURSOR (APOLIPOPROTEIN H) (APO-H) (ACTIVATED
 DE PROTEIN C-BINDING PROTEIN) (APC INHIBITOR) (B2GPI).
 GN APOH OR B2GPI.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 92372000.
 RA Nonaka M., Matsuda Y., Shirosaki T., Moriwak K., Natsune-Sakai S.;
 RT "Molecular cloning of mouse beta 2-glycoprotein I and mapping of the
 RT gene to chromosome 11.";
 RN Genomics 13:1082-1087(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CBA/J; TISSUE-LIVER;
 RX MEDLINE; 94242017.
 RA Sellar G.C., Steel D.M., Zafiroopoulos A., Seery L.T.,
 RA Whitehead A.S.;
 RT "Characterization, expression and evolution of mouse beta 2-
 RT glycoprotein I (apolipoprotein H).";
 RN Biochem. Biophys. Res. Commun. 200:1521-1528(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE-LIVER;
 RA Kristensen T.;
 RT "Structure of the human beta-2-glycoprotein I gene.";
 CC Submitted (FEH-1997) to the EMBL/Genbank/DBJ databases.
 CC -I- FUNCTION: BINDS TO VARIOUS KINDS OF NEGATIVELY CHARGED SUBSTANCES
 CC SUCH AS HEPARIN, PHOSPHOLIPIDS, AND DEXTRAN SULFATE. MAY PREVENT
 CC ACTIVATION OF THE INTRINSIC BLOOD COAGULATION CASCADE BY BINDING
 CC TO PHOSPHOLIPIDS ON THE SURFACE OF DAMAGED CELLS.
 CC -I- TISSUE SPECIFICITY: PLASMA.
 CC -I- SIMILARITY: CONTAINS 4 SUSHI (SCR) REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D10056; BAA00945.1; -;
 DR EMBL; S70439; AAB30789.1; -;
 DR EMBL; Y11356; CAA72190.1; -;

FT	DISULFID	549	549	INTERCHAIN (WITH BETA CHAIN) (POTENTIAL).
FT	DISULFID	561	561	INTERCHAIN (WITH BETA CHAIN) (POTENTIAL).
FT	CARBOHYD	66	66	POTENTIAL.
FT	CARBOHYD	221	221	POTENTIAL.
FT	CARBOHYD	525	525	POTENTIAL.
FT	CARBOHYD	602	602	POTENTIAL.
SO	SEQUENCE	610 AA:	68886 MW: D806B270E8A06B58 CRC64:	
Query Match				
Best Local Similarity		14.1%;	Score 357;	DB 1: Length 610;
Matches		85; Conservative	71; Mismatches 131;	Indels 30; Gaps 25
Db	66	NENREFETGLTAYTCRPPRYRISRRKRLIDGDGMV-KYKEF--CYKKRCENPGELLNGQ	122	
QY	20	SDQTEPGEGQALYKCRPGGR-SLGVWIMVCR-KGEWALNPFLKCKCRKCGHGPSTPGGT	77	
Db	123	VIYKT-D-YSFSEIEFSCSEGYVLISANSYCOLDRGYVSDPLPQCIACKEPPTI	180	
QY	78	FTLTGNNVEYGVKAVYVNCNEDYQLLEGIN-YRCDPIDG--WTNDIPICEYVKKCLPVTAP	134	
Db	181	SNGR-HNGG-DEDF-VYTGSSVYVSCDRDSMLGKASISCRVYNKTIQVWSSPSPCKKV	237	
QY	135	ENKIVSSAMEPDREHFHFGQAVRFVNCNGYKIEDEDMHCS-DD--GWSKEKPKCWEI	190	
Db	238	ICQPVKXGKGTSGGPIYTYQOSIYVACKGFRLGSLDILHCEDNSMNPPTCELN	297	
QY	191	SCKSPVDNPSPISQ-KIITYKENEFQYCKMGWEGYSEKGDVNC-TESGRPLP-SCEEK	247	
Db	298	GCGLPHIPHALMERYDHQTEQGVYDIGFVLSYKCHFQKPEPTDPTTVYQCSNLEWS	357	
QY	248	SC-DNPIPIPG--D-YF-PLRKHTGD-E-I-TYCRNGFYATRGNTA-KCTST-GWI	297	
Db	358	PYECKEVCOCPEPNLNN	374	
QY	298	PAPRCULKPCDYPDIKH	314	
RESULT 13				
ID	LEM3_MOUSE	STANDARD:	PRT:	768 AA.
AC	Q01102;			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	P-SELECTIN PRECURSOR (GRANULE MEMBRANE PROTEIN 140) (GMP-140) (PADGEM)			
DE	(CP622) (LEUKOCYTE-ENDOTHELIAL CELL ADHESION MOLECULE 3) (LECAM5).			
GN	SELP OR GRMP.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
OC	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
RN	[1]			
RX	SEQUENCE FROM N.A.			
RX	MEDLINE: 92340571.			
RA	Weller A., Isenmann S., Vestweber D.;			
RA	"Cloning of the mouse endothelial selectins. Expression of both E-			
RA	and P-selectin is inducible by tumor necrosis factor alpha."			
RL	J. Biol. Chem. 267:15176-15183(1992).			
RN	[2]			
RN	SEQUENCE FROM N.A.			
RX	MEDLINE: 92345617.			
RA	Sanders W.E. Jr., Wilson R.W., Ballantyne C.M., Baudet A.L.;			
RA	"Molecular cloning and analysis of in vivo expression of murine P-			
RL	selectin."			
RL	Blood 80:795-800(1992).			
CC	-I- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS			
CC	TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE			
CC	INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH			
CC	LEUKOCYTES. THE LIGAND RECOGNIZED IS SIALYL-LEWIS X.			
CC	-I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	-I- TISSUE SPECIFICITY: STORED IN THE ALPHA-GRANULES OF PLATELETS			
CC	AND WEIBEL-PALADE BODIES OF ENDOTHELIAL CELLS. UPON CELL			
CC	ACTIVATION BY AGONISTS, P-SELECTIN IS TRANSPORTED RAPIDLY TO			
CC	THE CELL SURFACE.			
CC	-I- INDUCTION: BY TUMOR NECROSIS FACTOR ALPHA.			

FT DISULFID 364 389 BY SIMILARITY.
 FT DISULFID 396 439 BY SIMILARITY.
 FT DISULFID 425 450 BY SIMILARITY.
 FT DISULFID 454 505 BY SIMILARITY.
 FT DISULFID 486 515 BY SIMILARITY.
 FT DISULFID 524 567 BY SIMILARITY.
 FT DISULFID 553 578 BY SIMILARITY.
 FT DISULFID 582 636 BY SIMILARITY.
 FT DISULFID 616 646 BY SIMILARITY.
 FT CARBOHYD 162 162 POTENTIAL.
 FT CARBOHYD 545 545 CELL ATTACHMENT SITE.
 FT SITE 617 619 C-> F (IN F13B DEFICIENCY).
 FT VARIANT 450 450 /F13B-VAR_007475.
 SQ SEQUENCE 661 AA; 75491 MW; 57A2FB46560857F2 CRC64;

Query Match 13.4%; Score 340; DB 1; Length 661;
 Best Local Similarity 28.7%; Pred. No. 6.15e-61;
 Matches 89; Conservative 57; Mismatches 134; Indels 30; Gaps 24;

Db 178 YECATGYTTAGKRTKEVECLTYGN-SLTP--KCTKLKSSLRLENGYFH-PVKQYEE 233
 QY 32 YKCRPGY-RSLGN-VIMV-CRKGEMVALNPLKCGKRCGHGDPFGTTLTGAVPEY 88
 Db 234 GDVVOFCHENYVLSGS-DLIQCTNFGWPPSPVCEGRNRNCPPEPLPINSKIOTHT-T 291
 QY 89 GVKAVYTCNEGYOLLGEINRECDTGDWTNDIPICEV--VKCLPYTAPENGIIVSSAMEP 146
 Db 292 YR--H-GSIHIECELNFEIHGSAEIRC-EDGKWT-EPPKCIEGQEKVACEEPPFIENGA 346
 QY 147 DREYHFGQAVAFVYCNISGKTIKDEDEMHCSDDGFMSKEPKCFE----ISCKSPDYI-NGS 201
 Db 347 ANHSHKIYNGDKVTYACKSGYLLHGSNEITCNRGKWTLPPECVENNENCKHPPVWNGA 406
 QY 202 PISOKIITYKENERPQYCKNMGYEYSEKGDVCTESGWRPLPSC-EER-SCDNP-YIPNGD 258
 Db 407 VADGIILASVATGSSVEYRC-NEYV-LLRGSKISRCEQKWSPPVC-LEPCT-VNVDYMN 462
 QY 259 YSP-LRIKHRTGDEITYQCRNGFYPATRGNT-AKCTSTGWLIPARCTLPCDYPDIKHGG 316
 Db 463 RNNIEMKWKY 472
 QY 317 LYHENMRRPY 326

Search completed: Thu Jun 8 21:46:19 2000
 Job time : 15 secs.

MIPS RELEASE

(TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mpsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jun 8 21:46:38 2000. Maspar time 30.28 Seconds

Tabular output not generated. 753.376 Million cell updates/sec

Title: >US-09-316-163-11
Description: (1-329) ~~from~~ US09316163.pep
Perfect Score: 2539
Sequence: 1 EDONELPPRRNTETILIGSNS.....PDIKHGGLYHNNRRPFV 329

Scoring table: PAM 150
Gap 11

Searched: 228878 seqs, 80334122 residues

Post-processing: Minimum Match 08
Listing first 45 summaries

Database:

sptrembl12
1:sp.archaea 2:sp.bacteria 3:sp.fungi 4:sp.human
5:sp.invertebrate 6:sp.mammal 7:sp.mmc 8:sp.organelle
9:sp.phage 10:sp.plant 11:sp.rodent 12:sp.unclassified
13:sp.vertebrate 14:sp.virus

Statistics: Mean 43.988; Variance 63.846; scale 0.689

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	2539	100.0	449	4	Q14570	COMPLEMENT FACTOR H PR	0.00e+00
2	1236	48.7	669	6	Q28085	CCP MODULES 3-12, WITH	0.00e+00
3	603	23.7	1053	13	Q91275	COMPLEMENT REGULATOR	1.99e-130
4	476	18.7	657	4	Q14006	COMPLEMENT H FACTOR (F	1.99e-95
5	438	17.3	452	11	Q61407	COMPLEMENT FACTOR H-RE	4.10e-85
6	436	17.2	645	14	Q9WR2	COMPLEMENT BINDING PRO	1.42e-84
7	433	17.1	303	11	Q61405	COMPLEMENT FACTOR H-RE	9.20e-84
8	429	16.9	550	14	Q40912	ORF 4.	1.10e-82
9	428	16.9	550	14	P88903	ORF 4.	2.05e-82
10	426	16.8	522	6	Q28769	COMPLEMENT RECEPTOR (F	7.11e-82
11	423	16.7	2014	6	Q29530	COMPLEMENT RECEPTOR 1	4.57e-81
12	412	16.2	1911	6	Q29528	COMPLEMENT RECEPTOR 1	4.14e-78
13	409	16.1	2039	4	Q16745	COMPLEMENT RECEPTOR 1	2.64e-77
14	409	16.1	2489	4	Q16744	COMPLEMENT RECEPTOR 1	2.64e-77
15	403	15.9	661	6	Q29531	COMPLEMENT RECEPTOR 1	1.07e-75
16	398	15.7	259	14	P87616	41KBP FRAGMENT FROM LE	2.33e-74
17	389	15.3	522	6	Q28797	UNKNOWN PROTEIN (FRAGM	5.90e-72
18	388	15.3	560	5	Q22328	COSMID T07H6.	1.09e-71
19	382	15.0	360	14	Q8XT08	COMPLEMENT CONTROL PRO	4.32e-70
20	378	14.9	417	11	Q35520	512 ANTIGEN (FRAGMENT)	5.00e-69

ALIGNMENTS

RESULT ID	Query Match	Score	DB	Length	Query Match	Score	DB	Length
1	Q14570	100.0%	449	449	Q14570	100.0%	449	449
2	Q14570	100.0%	449	449	Q14570	100.0%	449	449
3	Q14570	100.0%	449	449	Q14570	100.0%	449	449
4	Q14570	100.0%	449	449	Q14570	100.0%	449	449
5	Q14570	100.0%	449	449	Q14570	100.0%	449	449
6	Q14570	100.0%	449	449	Q14570	100.0%	449	449
7	Q14570	100.0%	449	449	Q14570	100.0%	449	449
8	Q14570	100.0%	449	449	Q14570	100.0%	449	449
9	Q14570	100.0%	449	449	Q14570	100.0%	449	449
10	Q14570	100.0%	449	449	Q14570	100.0%	449	449
11	Q14570	100.0%	449	449	Q14570	100.0%	449	449
12	Q14570	100.0%	449	449	Q14570	100.0%	449	449
13	Q14570	100.0%	449	449	Q14570	100.0%	449	449
14	Q14570	100.0%	449	449	Q14570	100.0%	449	449
15	Q14570	100.0%	449	449	Q14570	100.0%	449	449
16	Q14570	100.0%	449	449	Q14570	100.0%	449	449
17	Q14570	100.0%	449	449	Q14570	100.0%	449	449
18	Q14570	100.0%	449	449	Q14570	100.0%	449	449
19	Q14570	100.0%	449	449	Q14570	100.0%	449	449
20	Q14570	100.0%	449	449	Q14570	100.0%	449	449
21	Q14570	100.0%	449	449	Q14570	100.0%	449	449
22	Q14570	100.0%	449	449	Q14570	100.0%	449	449
23	Q14570	100.0%	449	449	Q14570	100.0%	449	449
24	Q14570	100.0%	449	449	Q14570	100.0%	449	449
25	Q14570	100.0%	449	449	Q14570	100.0%	449	449
26	Q14570	100.0%	449	449	Q14570	100.0%	449	449
27	Q14570	100.0%	449	449	Q14570	100.0%	449	449
28	Q14570	100.0%	449	449	Q14570	100.0%	449	449
29	Q14570	100.0%	449	449	Q14570	100.0%	449	449
30	Q14570	100.0%	449	449	Q14570	100.0%	449	449
31	Q14570	100.0%	449	449	Q14570	100.0%	449	449
32	Q14570	100.0%	449	449	Q14570	100.0%	449	449
33	Q14570	100.0%	449	449	Q14570	100.0%	449	449
34	Q14570	100.0%	449	449	Q14570	100.0%	449	449
35	Q14570	100.0%	449	449	Q14570	100.0%	449	449
36	Q14570	100.0%	449	449	Q14570	100.0%	449	449
37	Q14570	100.0%	449	449	Q14570	100.0%	449	449
38	Q14570	100.0%	449	449	Q14570	100.0%	449	449
39	Q14570	100.0%	449	449	Q14570	100.0%	449	449
40	Q14570	100.0%	449	449	Q14570	100.0%	449	449
41	Q14570	100.0%	449	449	Q14570	100.0%	449	449
42	Q14570	100.0%	449	449	Q14570	100.0%	449	449
43	Q14570	100.0%	449	449	Q14570	100.0%	449	449
44	Q14570	100.0%	449	449	Q14570	100.0%	449	449
45	Q14570	100.0%	449	449	Q14570	100.0%	449	449

Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 19 EDCEMLPPRRMTTEITGSMQOTPEGTOAIYKCPGRRSGNYIMVCRKGEWALNLR 78
QY 1 EDCEMLPPRRMTTEITGSMQOTPEGTOAIYKCPGRRSGNYIMVCRKGEWALNLR 60
Db 79 KQKRCPCGHPDTPFGFTTLTGAVNEFYGVAVYTCNGYOLLGEINREDDTGWTNDI 138
QY 61 KQKRCPCGHPDTPFGFTTLTGAVNEFYGVAVYTCNGYOLLGEINREDDTGWTNDI 120
Db 139 PICCVVCLPTAPENKGVSSAEPREYHFGQAVRVVCSGKIKESDEMHCSDGFW 198
QY 121 PICCVVCLPTAPENKGVSSAEPREYHFGQAVRVVCSGKIKESDEMHCSDGFW 180
Db 199 SKERPKCVESICKSPDYNGSPISOKIITYKNERFQYCKNNGYSESGDAVCTESGWRP 258
QY 181 SKERPKCVESICKSPDYNGSPISOKIITYKNERFQYCKNNGYSESGDAVCTESGWRP 240
Db 259 LPSCSEKSCNPNYPNDYSPLRKHKRTGDEITYQCNGEYPATRGWTAKCTGTGWTAP 318
QY 241 LPSCSEKSCNPNYPNDYSPLRKHKRTGDEITYQCNGEYPATRGWTAKCTGTGWTAP 300
Db 319 RCTLKPCDYPDIKGGIYHNMRRPYPV 347
QY 301 RCTLKPCDYPDIKGGIYHNMRRPYPV 329

RESULT 2
ID Q28085 PRELIMINARY; PRT; 669 AA.
AC Q28085;

DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBLrel. 12, Last annotation update)
DE CCP MODULES 3-12, WITH PARTS OF CCP 2 AND 13 (FRAGMENT).
OS Bos taurus (Bovine).
RT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;
RN Bovinae; Bos.

RP SEQUENCE FROM N.A.
RC TISSUE-LIVER;
RA MEDLINE; 96202005.
RA SOAMES C.J., DAY A.J., SIM R.B.;
RT "Prediction from sequence comparisons of residues of factor H involved
in the interaction with complement component C3b."
RL Biochem. J. 315:523-531(1996).
DR EMBL; X98697; CA67257.1; -.
DR HSSP; F10998; IYVC.
DR PFAM; PF00084; sushi; 11.
FT NON_TER 1
FT SEQUENCE 669 AA; 75683 MW; FAF0D174 CRC32;

Query Match 48.7%; Score 1236; DB 6; Length 669;
Best Local Similarity 61.4%; Pred. No. 0.00e+00;
Matches 153; Conservative 35; Mismatches 59; Indels 2; Gaps 1;

Db 7 LAEGQFEYGAIVYTCDEGYOMVGMNREDDTNDICEVVKCLPTPEPNCGI 66
QY 80 LTGQVFEYGAIVYTCDEGYOMVGMNREDDTNDICEVVKCLPTPEPNCGI 139
Db 67 FSDALEPDEYTYGVVFECSNGYMLDPRKQIHCSAGSAGVSAETPKVEICKSPVYIN 126
QY 140 VSSAEPDREYHFGQAVRVVCSGKIKESDEMHCSDGFWKPKVEICKSPDYIN 199
Db 127 GOAVLPKATYONENVOYRCAGFEYGGQDVTCTKSGTTPATCIETTCDDPRIPNGY 186
QY 200 GSPISOKIITYKNERFQYCKNNGYSESGDAVCTESGWRPPLPSCSEKSCNPNYP 259
Db 187 RPELSKYRGQKRIYCKKGFEPRTGATDGTGAWPVVPCAKPKSPYIKRGLY 246
QY 260 SPLRKHTGTGEITYQCNGEYPATRGWTAKCTGTGWTAPRCTLKPCDYPDIKGGIYH 319

Db 247 SY--RGYFP 253
QY 320 ENMRPYPF 328

RESULT 3
ID Q91275 PRELIMINARY; PRT; 1053 AA.
AC Q91275;

DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBLrel. 12, Last annotation update)
DE COMPLEMENT REGULATORY PLASMA PROTEIN.
OS Parabrax nebulifer (barred sand bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopeterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
OC Perciformes; Percoidae; Serranidae; Parabrax.
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE-LIVER;
RA MEDLINE; 94318039.
RA DAHMEN A., KAIJOH T., ZIPPEL P.F., GIGLI I.;
RT "Cloning and characterization of a cDNA representing a putative
complement-regulatory plasma protein from barred sand bass (Parabrax
nebulifer)."
RL Biochem. J. 301:391-397(1994).
DR EMBL; L21703; AAA92556.1; -.
DR HSSP; P08603; IHPH.
DR PFAM; PF00084; sushi; 16.
FT SEQUENCE 1053 AA; 117597 MW; 0DF68EDB CRC32;

Query Match 23.7%; Score 603; DB 13; Length 1053;
Best Local Similarity 34.1%; Pred. No. 1.99e-130;
Matches 101; Conservative 49; Mismatches 129; Indels 17; Gaps 13;

Db 47 EASYPCGROVRYGCVNYG-S-GFFKLVCYEGKWEYRGA--KCQPRSCGHPDAOFADPHL 102
QY 21 DQTYEGTQAIYKCRGYSRLGNVIMVCRKGEWALNLRKQKRCGHPDTPFGFTL 80
Db 103 AEGNDFVGSKYVYTCOKGYOMVSRINRYRCVLAEGMDGVVPCESQCC-PLIHDNNVOY 161
QY 81 TGGNVFEYGVAVYTCNGYOLLGEINREDDTGWTNDICEVVKCLPTAPENKGIY 140
Db 162 ICG--PE-EATFGVNVREFSCRSSEILDSPELYCDEMGAPPKCAIKALPIEN 218
QY 141 SSAFEPRREYHFGQAVRVVCSGKIKESDEMHCSDGFWKPKVEICKSPDYIN 199
Db 219 GNVPAIRKRYKNDVLAHECDRAFKHIDRPSTCIKQIKAEWSPPLCEISIKRLTMDG 278
QY 200 GSPISOKIITYKNERFQYCKNNGYSESGDA-V--CTESGWRPPLPSCSEKSCNPNYP 256
Db 279 TRYEPAYRNLFSPGELTKLVICARTSWISTPOETSVVTCQDNGEISIRPTQGVNRC 334
QY 257 GDYSPL-RHKRTGDEITYQC-RNGFYPATRGNT--AKCTSTG-WTPAPRCTLKPC 307

RESULT 4
ID Q14006 PRELIMINARY; PRT; 657 AA.
AC Q14006;

DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBLrel. 12, Last annotation update)
DE COMPLEMENT H FACTOR (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominiidae; Homo.

RP SEQUENCE FROM N.A.
RA MEDLINE; 88025472.
RA DAY A.J., RIPOCHE J., LYONS A., MCINTOSH B., HARRIS T.J., SIM R.B.;
RT "Sequence analysis of a cDNA clone encoding the C-terminal end of
human complement factor H.";

BL Bioscl. Rep. 7:201-207(1987).
 DR EMBL; M17517; AAA52016.1; -
 DR HSSP; P08603; 1HFH.
 DR PFAM; PF00084; sushi; 11.
 FT NON_TER 1
 SO SEQUENCE 657 AA; 74247 MW; F4AB5238 CRC32;

Query Match 18.7%; Score 476; DB 4; Length 657;
 Best Local Similarity 27.5%; Pred. No. 1,99e-95;
 Matches 86; Conservative 72; Mismatches 132; Indels 23; Gaps 22;

Db 296 CSQ-PROTEINSSSSSSQSSAHGKLTSTCEGFRISENETTCYMGK-S-SP-PQ 351
 QY 3 CNEPLPRNTEILGWSWQ-TYPEGQALTKCRPGRLSLGNVIMVCRKGEVALNLRK 61
 Db 352 CEGLPCKRPEISHGVNAMS-D-YQGEETVYKCFEGGIDGPAIAK-CLGEKMSHP-P 408
 QY 62 CCKRPGHGTPTGTTLTGNGVFEYGVKAVYTCNNEGTLGELINRECDGTNDIP 121
 Db 409 SCIKTDLSPFENA-IPGEKK-D-YKAGEOYTYCAIYKMDASNTYCINSR-WT 464
 QY 122 ICEVVKCLPVTAPENGKIVSSAMEPDEYHFGQAVRFVCSNGYKIEGDEHSCDDGFW 181
 Db 465 -GRPCRTSCVNPPTVQNAIVSRQMSKYPSEGRVRYQCSPEMFGD-EVWACLGNW 522
 QY 182 KEKPCVEISC-KSPDIVNGSPIS-QKIIYKENERFOYCKMNGE-YSERGDVCTESGW 238
 Db 523 TEPPOCKDSTGCGPPPIDNGITFPLSYAPASSVYEOCON-LYOLEGNKRITCRNG 581
 QY 239 RLPLSCERKS--CDNPY-IPNGDYSPLRK-HRIGDELTYCCRNNGFYATGNTAKCTST 294
 Db 582 QWSEPPKC-LHPC 593
 QY 295 GWIPARCTLKPC 307

RESULT 5
 ID 061407 PRELIMINARY; PRT; 452 AA.
 AC 061407;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE COMPLEMENT FACTOR H-RELATED PROTEIN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RP SEQUENCE FROM N.A.
 RA VIK D.P., MUNOZ-CANOVES P., KOZONO H., MARTIN L.G., TACK B.F.,
 RA CHAPLIN D.D.;
 RL J. Biol. Chem. 0:0-0(0).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA VIK D.P., MUNOZ-CANOVES P., KOZONO H., MARTIN L.G., TACK B.F.,
 RA CHAPLIN D.D.;
 RL Submitted (Apr-1990) to the EMBL/GenBank/DBJ databases.
 DR EMBL; M29010; AAA37415.1; -
 DR HSSP; P08603; 1HFI.
 DR PFAM; PF00084; sushi; 7.
 SO SEQUENCE 452 AA; 51602 MW; 45C0B61 CRC32;

Query Match 17.3%; Score 438; DB 11; Length 452;
 Best Local Similarity 56.4%; Pred. No. 4.10e-85;
 Matches 53; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

Db 15 TAWLSTAKEETKSPYILNGITPRIRIHSRDEIRYECNFGFYPTGTSKCTPTG 74
 QY 236 SGWRPLPSCERKSCDNPYIPNGDYSPLRIKHRTGDELTYCCRNNGFYATGNTAKCTSTG 295
 Db 75 WIPVRCGLKPCPEPQKRYGLYEEESLRPNFV 108
 QY 296 WIPARCTLKPCDYPDKHGLYHENMRPRYFV 329

RESULT 6
 ID 09WRU2 PRELIMINARY; PRT; 645 AA.
 AC 09WRU2;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE COMPLEMENT BINDING PROTEIN.
 OS Macaca mulatta rhadinovirus 17577.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 99174001.
 RA SEARLES R.P., BERGQUAM E.P., AXTHELM M.K., WONG S.W.;
 RT "Sequence and genomic analysis of a Rhesus macaque rhadinovirus with
 RT similarity to Kaposi's sarcoma-associated herpesvirus/human
 RL herpesvirus 8.";
 RL J. Virol. 73:3040-3053(1999).
 DR EMBL; AF083501; AAD21332.1; -
 SO SEQUENCE 645 AA; 71526 MW; 60FB82D6 CRC32;

Query Match 17.2%; Score 436; DB 14; Length 645;
 Best Local Similarity 30.9%; Pred. No. 1.42e-84;
 Matches 72; Conservative 50; Mismatches 94; Indels 17; Gaps 14;

Db 337 EKYSVGSVELICRPGFTKQSTVSECLSNGTWAPNA--KCHRRKCPDPELLNGEYI 394
 QY 22 QITPEGQALTKCRPGRLSLGNVIM-C-RKGEVALNLRKQCRKPGDPTGTPT 79
 Db 395 VTSGEDAFKYGITNYKCNQGYLLGSMWRICMLKDLDTVMEKAPICDIEKCPQ 454
 QY 80 LTGG-NVEYGVKAVYTCNNEGTLGELIN-RECDGTNDIPCEVVKCLPVTA 133
 Db 455 ITNCKY--HEVYK-D-YQYIDTYTFSCNRDSIVGDMETTCISNT--MKNPFCQDITCS 509
 QY 134 PENKIVSAMEPDEYHFGQAVRFVCSNGYKIEGDEHSCDDGFWSEKPCVVISCK 193
 Db 510 APNIAHGKLTGSSSVYKYGQSVTIGCTGFTLIGSEISTCKDSWDPPLPTC 562
 QY 194 SPDVINGSPIS-QKIIYKENERFOYCKMNGE-YSERGDVCTESGW-RLPLSC 244

RESULT 7
 ID 061405 PRELIMINARY; PRT; 303 AA.
 AC 061405;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE COMPLEMENT FACTOR H-RELATED PROTEIN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA VIK D.P., MUNOZ-CANOVES P., KOZONO H., MARTIN L.G., TACK B.F.,
 RX MEDLINE; 90153969.
 RA VIK D.P., MUNOZ-CANOVES P., KOZONO H., MARTIN L.G., TACK B.F.,
 RA CHAPLIN D.D.;
 RL "Identification and sequence analysis of four complement factor H-
 RL related transcripts in mouse liver.";
 RL J. Biol. Chem. 265:3193-3201(1990).
 DR EMBL; M29007; AAA37413.1; -
 DR HSSP; P10998; 1VVC.
 DR PFAM; PF00084; sushi; 4.
 SO SEQUENCE 303 AA; 34498 MW; 39350FD1 CRC32;

Query Match 17.1%; Score 433; DB 11; Length 303;
 Best Local Similarity 55.3%; Pred. No. 9.20e-84;
 Matches 52; Conservative 17; Mismatches 25; Indels 0; Gaps 0;

Db 15 TAWLSTAKEETKSPYILNGITPRIRIHSRDEIRYECNFGFYPTGTSKCTPTG 74
 QY 236 SGWRPLPSCERKSCDNPYIPNGDYSPLRIKHRTGDELTYCCRNNGFYATGNTAKCTSTG 295

DR EMBL: L17394; AAB60694.1; JOINED.
 DR EMBL: L17395; AAB60694.1; JOINED.
 DR EMBL: L17396; AAB60694.1; JOINED.
 DR EMBL: L17397; AAB60694.1; JOINED.
 DR EMBL: L17398; AAB60694.1; JOINED.
 DR EMBL: L17399; AAB60694.1; JOINED.
 DR EMBL: L17400; AAB60694.1; JOINED.
 DR EMBL: L17401; AAB60694.1; JOINED.
 DR EMBL: L17402; AAB60694.1; JOINED.
 DR EMBL: L17403; AAB60694.1; JOINED.
 DR EMBL: L17404; AAB60694.1; JOINED.
 DR EMBL: L17405; AAB60694.1; JOINED.
 DR EMBL: L17406; AAB60694.1; JOINED.
 DR EMBL: L17407; AAB60694.1; JOINED.
 DR EMBL: L17408; AAB60694.1; JOINED.
 DR EMBL: L17409; AAB60694.1; JOINED.
 DR EMBL: L17410; AAB60694.1; JOINED.
 DR EMBL: L17411; AAB60694.1; JOINED.
 DR EMBL: L17412; AAB60694.1; JOINED.
 DR EMBL: L17413; AAB60694.1; JOINED.
 DR EMBL: L17414; AAB60694.1; JOINED.
 DR EMBL: L17415; AAB60694.1; JOINED.
 DR EMBL: L17416; AAB60694.1; JOINED.
 DR EMBL: L17417; AAB60694.1; JOINED.
 DR EMBL: L17418; AAB60694.1; JOINED.
 DR EMBL: L17419; AAB60694.1; JOINED.
 DR EMBL: L17420; AAB60694.1; JOINED.
 DR EMBL: L17421; AAB60694.1; JOINED.
 DR EMBL: L17422; AAB60694.1; JOINED.
 DR EMBL: L17423; AAB60694.1; JOINED.
 DR HSSP: P08603; IHFI.
 DR PFAM: PF00084; sush1; 30.
 DR SEQUENCE 2039 AA; 223603 MW; A4972215 CRC32.

Query Match 16.18; Score 409; DB 4; Length 2039;
 Best Local Similarity 31.48; Pred. No. 2,64e-77;
 Matches 97; Conservative 66; Mismatches 110; Indels 36; Gaps 30;

DB 63 PFIGTYLNECRPGYSGRPFST-ICLNKSVWTKAKD-R-CRRKSCRRPPDPVNGMAYVIR 119
 QY 24 YPEGTQAIYKCRPGYRSLGNVIMWC-RKGEWVALNLRKQCRPGCHGDTPTGTFTLIG 82
 DB 120 G-I-OFGSQIKYCTKGRYLIGSSSSTATCIISGDTVINDNETPICDRIPCGLPPTIT-NGD 176
 QY 83 GNVFEYGVKAVYTCNMGYOLGGE--INPREG-DTIDGTNDIPICEVYKC-LPVYAPENK 138
 DB 177 FISTNRE-N--FHYGSVYTYRCNPGSGKRVFELVEPSIYCTSDNDQVIGWGPAPQCI 233
 QY 139 IVSSAMEPDRHYHFGQAVRFVNCNGY--KI-E--GDEEMHC-S-DD--GFWSKKEPKCV 188
 DB 234 IPNKCTPPNVEGILVSDNLSFLSNEVEYFRQCPGFVYKGPFRVYKQALNKKEPELPSC 293
 QY 189 EIS-CKSPDIVNGSPISQ-KIITYKENRFOYKCMNGEYSESGDAVCTE-SGWRP-LPSC 244
 DB 294 S-RVCOOPRDLVHAARTORDKNDFSPGOEVYSCPEG-Y-DIRGAASMRCTPPQGDWSPAA 350
 QY 245 EEKSCDND-YIPNGYSPL-RIKHRTGDEITYQCRNGFYPAIRGNTA-KCISTG-WIPA- 299
 DB 351 PTCVKSQCD 359
 QY 300 PCTLKPCD 308

RESULT 14
 ID Q16744; PRELIMINARY; PRT; 2489 AA.
 AC Q16744;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
 DE COMPLEMENT RECEPTOR 1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 RC [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE; 94065173.
 RA VIK D.P., WONG W.W.;
 RT Structure of the gene for the F allele of complement receptor type 1
 and sequence of the coding region unique to the S allele.*;
 RL J. Immunol. 151:6214-6224(1993).
 RN [2]

RP SEQUENCE FROM N.A.

RA VIK D.P., WONG W.W.;
 RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.

DR EMBL: L17418; AAB60695.1; JOINED.
 DR EMBL: L17390; AAB60695.1; JOINED.
 DR EMBL: L17391; AAB60695.1; JOINED.
 DR EMBL: L17392; AAB60695.1; JOINED.
 DR EMBL: L17393; AAB60695.1; JOINED.
 DR EMBL: L17394; AAB60695.1; JOINED.
 DR EMBL: L17395; AAB60695.1; JOINED.
 DR EMBL: L17396; AAB60695.1; JOINED.
 DR EMBL: L17397; AAB60695.1; JOINED.
 DR EMBL: L17398; AAB60695.1; JOINED.
 DR EMBL: L17399; AAB60695.1; JOINED.
 DR EMBL: L17400; AAB60695.1; JOINED.
 DR EMBL: L17401; AAB60695.1; JOINED.
 DR EMBL: L17402; AAB60695.1; JOINED.
 DR EMBL: L17403; AAB60695.1; JOINED.
 DR EMBL: L17404; AAB60695.1; JOINED.
 DR EMBL: L17405; AAB60695.1; JOINED.
 DR EMBL: L17406; AAB60695.1; JOINED.
 DR EMBL: L17407; AAB60695.1; JOINED.
 DR EMBL: L17408; AAB60695.1; JOINED.
 DR EMBL: L17409; AAB60695.1; JOINED.
 DR EMBL: L17410; AAB60695.1; JOINED.
 DR EMBL: L17411; AAB60695.1; JOINED.
 DR EMBL: L17412; AAB60695.1; JOINED.
 DR EMBL: L17413; AAB60695.1; JOINED.
 DR EMBL: L17414; AAB60695.1; JOINED.
 DR EMBL: L17415; AAB60695.1; JOINED.
 DR EMBL: L17416; AAB60695.1; JOINED.
 DR EMBL: L17417; AAB60695.1; JOINED.
 DR EMBL: L17418; AAB60695.1; JOINED.
 DR EMBL: L17419; AAB60695.1; JOINED.
 DR EMBL: L17420; AAB60695.1; JOINED.
 DR EMBL: L17421; AAB60695.1; JOINED.
 DR EMBL: L17422; AAB60695.1; JOINED.
 DR EMBL: L17423; AAB60695.1; JOINED.
 DR EMBL: L17424; AAB60695.1; JOINED.
 DR EMBL: L17425; AAB60695.1; JOINED.
 DR EMBL: L17426; AAB60695.1; JOINED.
 DR EMBL: L17427; AAB60695.1; JOINED.
 DR EMBL: L17428; AAB60695.1; JOINED.
 DR EMBL: L17429; AAB60695.1; JOINED.
 DR EMBL: L17430; AAB60695.1; JOINED.
 DR HSSP: P08603; IHFI.
 DR PFAM: PF00084; sush1; 37.
 DR SEQUENCE 2489 AA; 272846 MW; 5869B6F9 CRC32;

Query Match 16.18; Score 409; DB 4; Length 2489;
 Best Local Similarity 31.48; Pred. No. 2,64e-77;

Matches 97; Conservative 66; Mismatches 110; Indels 36; Gaps 30;

DB 63 PFIGTYLNECRPGYSGRPFST-ICLNKSVWTKAKD-R-CRRKSCRRPPDPVNGMAYVIR 119
 QY 24 YPEGTQAIYKCRPGYRSLGNVIMWC-RKGEWVALNLRKQCRPGCHGDTPTGTFTLIG 82
 DB 120 G-I-OFGSQIKYCTKGRYLIGSSSSTATCIISGDTVINDNETPICDRIPCGLPPTIT-NGD 176
 QY 83 GNVFEYGVKAVYTCNMGYOLGGE--INPREG-DTIDGTNDIPICEVYKC-LPVYAPENK 138
 DB 177 FISTNRE-N--FHYGSVYTYRCNPGSGKRVFELVEPSIYCTSDNDQVIGWGPAPQCI 233
 QY 139 IVSSAMEPDRHYHFGQAVRFVNCNGY--KI-E--GDEEMHC-S-DD--GFWSKKEPKCV 188
 DB 234 IPNKCTPPNVEGILVSDNLSFLSNEVEYFRQCPGFVYKGPFRVYKQALNKKEPELPSC 293
 QY 189 EIS-CKSPDIVNGSPISQ-KIITYKENRFOYKCMNGEYSESGDAVCTE-SGWRP-LPSC 244

Db 294 S-RVQCAPPDVLHAERTORDNFSPOGEVYSCPEG-Y-DLRGAASMRCTPOGWSPPA 350
QY 245 EKSQCDNP-YIPNGDYSPL-RIKHRTGDEITYQCRNGFPATRGNTA-KCTSTG-WIPA- 299
Db 351 PTCEVKSQD 359
QY 300 PCTLKPCD 308

RESULT 15
ID 029531 PRELIMINARY; PRT; 661 AA.
AC 029531.
DT 01-NOV-1996 (TIREMBLrel. 01, Created)
DT 01-NOV-1996 (TIREMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TIREMBLrel. 12, Last annotation update)
DE COMPLEMENT RECEPTOR 1 PRECURSOR (FRAGMENT).
GN CRI.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Pan.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94292799.
RA BIRMINGHAM D.J., SHEN X.P., HOURCADE D., NICKELLS M.W., ATKINSON J.P.;
RT "Primary sequence of an alternatively spliced form of CRI. Candidate
RT for the 75,000 Mr complement receptor expressed on chimpanzee
RT erythrocytes."
RT J. Immunol. 153:691-700(1994).
DR EMBL: L24921; AAA51439.1; -.
DR HSSP: P10998; IYVC.
DR PIRAM, PR00084; susht; 9.
KW Signal; Alternative splicing.
FT NON_TER 1
FT SIGNAL <1 16 POTENTIAL.
FT CHAIN 17 >661 COMPLEMENT RECEPTOR 1.
FT NON_TER 661
SQ SEQUENCE 661 AA; 72966 MW; 9D78E262 CRC32;

Query Match

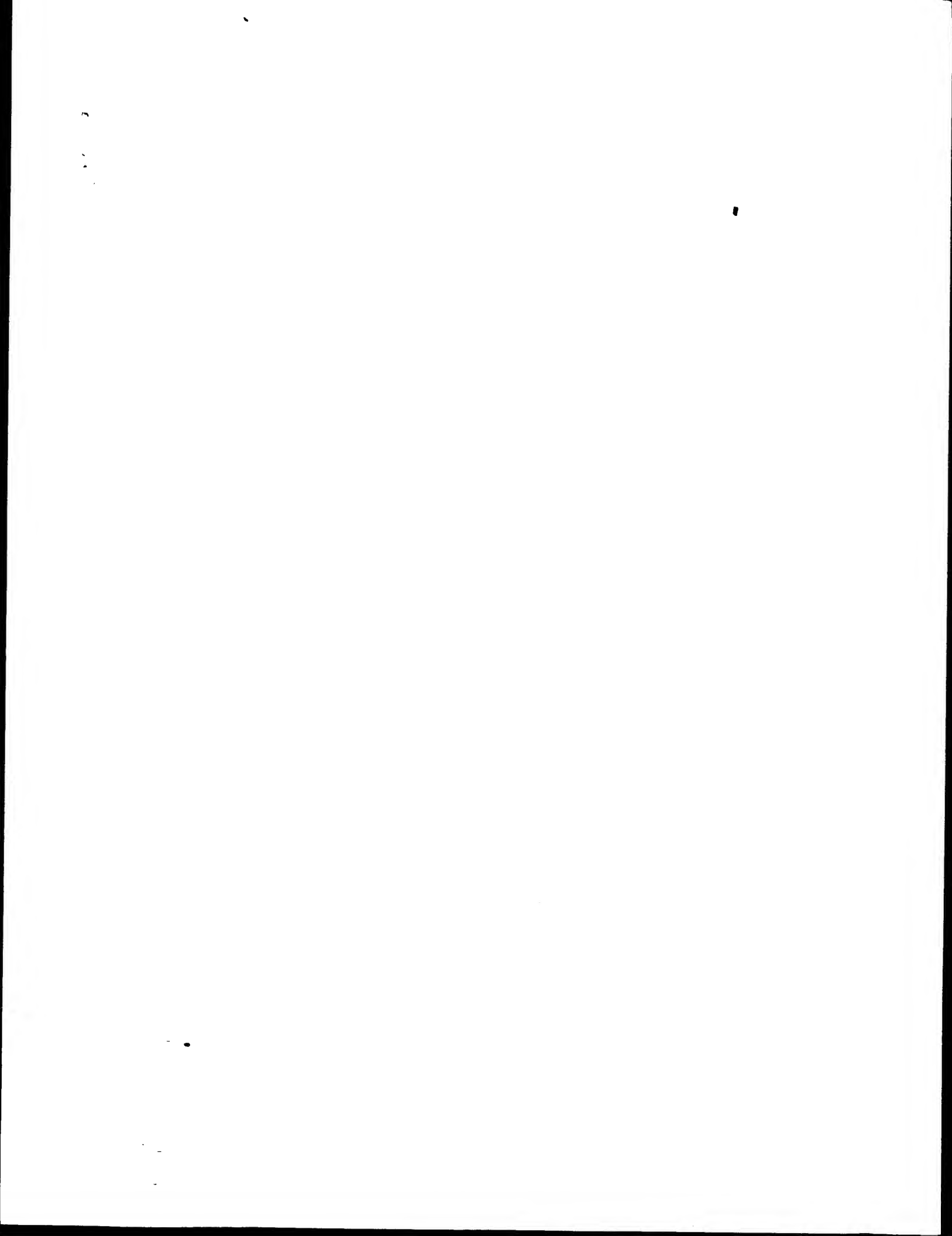
Best Local Similarity 15.98; Score 403; DB 6; Length 661;

Best Local Similarity 31.18; Pred. No. 1.07e-75;

Matches 96; Conservative 66; Mismatches 11; Indels 36; Gaps 30;

Db 38 FPIGTIANTCRPGYGRPSI-ICLNKSVWTKAD-R-CRRKSCRPPD-PYNGWVHY 93
QY 24 YPEGTOAIYICRPGYRSLGNVIMVC-RKGEVALNPLKCKQKRCGHPGDTPTGTTLTG 82
Db 94 KDI-QFGSQIKYSCCTKGYRLIGSSSATCIIISDPTIWMNETPICRIICGLPPTIT-NGD 151
QY 83 GNVEFVGKAVYTCNDEGYQLGE--INVRCC-DTGGWINDIPICEVYKC-LPYTAPENGK 138
Db 152 FISTNRE-N--FHVGSVYTYRCNPGSGGKRVFELVGEPSIYCTSNDDQVGIWSPAPOCI 208
QY 139 IVSASMEPDREYHFGQAVRFVCSNGY--KI-E--GDEEMHC-S-DD--GEWSKERPKCV 188
Db 209 IPNKCPTPNENGLVSDNLSFLSNVEVEFCOPGFVMKGPVVKCOALNKNPELPSC 268
QY 189 EIS-CKSPDIVNSPISQ-KIYKENRFYKCMGIEYERGDAYCTE-SGWRP-LPSC 244
Db 269 S-RVQCAPPDVLHAERTORDNFSPOGEVYSCPEG-Y-DLRGAASLRCTPOGWSPPAT 325
QY 245 EKSQCDNP-YIPNGDYSPL-RIKHRTGDEITYQCRNGFPATRGNTA-KCTSTG-WIPA- 299
Db 326 PTCEVKSQD 334
QY 300 PCTLKPCD 308

Search completed: Thu Jun 8 21:47:12 2000
Job time : 34 secs.



 MIPRELI (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mpsrch_dp protein - protein database search, using Smith-Waterman algorithm
 Run on: Thu Jun 8 21:49:10 2000: Maspar time 13.73 Seconds
 Tabular output not generated. 738.315 Million cell updates/sec

Title: >US-09-316-163-14
 Description: (1-428) from US09316163.pep
 Perfect Score: 3371
 Sequence: 1 EDCKGPPRENSILSGMS.....DYCYCENGSPPPKVRIRK 428

Scoring table: PAM 150
 Gap 11

Searched: 188963 seqs/23686106 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: a-gene:seq36
 1:gene:seq36

Statistics: Mean 34.511; Variance 134.258; scale 0.257

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1385	41.1	240	1 W39154	Human partial Complement	1.86e-134
2	1276	37.9	216	1 W39155	Clone PRB9PH410 CFH r	1.44e-122
3	712	21.1	578	1 Y09065	Human complement factor	1.36e-61
4	537	15.9	561	1 R13490	Human C4 binding prote	5.15e-43
5	445	13.2	543	1 R28557	CR1-4 (99H, 103E) anal	2.21e-33
6	445	13.2	543	1 R28552	CR1-4 (78T, 79D) analo	2.21e-33
7	446	13.2	720	1 W73147	Amino acid sequence of	1.74e-33
8	446	13.2	1930	1 W45899	Human complement recep	1.74e-33
9	446	13.2	2317	1 P92219	CR1 protein	1.74e-33
10	441	13.1	543	1 R28553	CR1-4 (85R, 87N) analo	5.77e-33
11	439	13.0	543	1 R28560	CR1-4 (114S) analogue.	9.31e-33
12	439	13.0	543	1 R28558	CR1-4 (109N, 110A, 111	9.31e-33
13	438	13.0	543	1 R28571	CR1-4 (364-367 NAAH) a	1.18e-32
14	437	13.0	543	1 R28550	CR1-4 (64K) analogue.	1.50e-32
15	439	13.0	1537	1 R11982	Partial human compleme	9.31e-33
16	439	13.0	2039	1 R36743	CR1.	9.31e-33
17	437	13.0	2039	1 R11810	Human complement type	1.50e-32
18	436	12.9	543	1 R28547	CR1-4 (52S, 53S, 54P)	1.91e-32
19	436	12.9	543	1 R28569	CR1-4 (36S-37E STRP)	1.91e-32
20	436	12.9	543	1 R28558	CR1-4 (347T, 349Y) ana	1.91e-32
21	435	12.9	543	1 R28565	CR1-4 (92T) analogue.	2.42e-32
22	434	12.9	543	1 R28565	CR1-4 (121Q) analogue.	3.08e-32
23	433	12.8	543	1 R28566	CR1-4 (318R, 319N) ana	3.91e-32

24	433	12.8	543	1 R28570	CR1-4 (266-274 KUKTOTN	3.91e-32
25	432	12.8	543	1 R28549	CR1-4 (64K, 65T) analo	4.96e-32
26	432	12.8	543	1 R28563	CR1-4 (117P) analogue.	4.96e-32
27	431	12.8	543	1 R28561	CR1-4 (115T) analogue.	6.30e-32
28	431	12.8	543	1 R28551	CR1-4 (65T) analogue.	6.30e-32
29	430	12.8	543	1 R28545	CR1-4 (37Y) analogue.	8.01e-32
30	430	12.8	543	1 R28562	CR1-4 (116K) analogue.	8.01e-32
31	428	12.7	543	1 R28544	CR1-4 (35E) analogue.	1.29e-31
32	427	12.7	543	1 R28548	CR1-4 (57K, 59K) analo	1.64e-31
33	426	12.6	543	1 R28564	CR1-4 (116K, 117P) ana	2.08e-31
34	425	12.6	543	1 R28567	CR1-4 (318-321 RNP) a	2.64e-31
35	424	12.6	543	1 R28556	CR1-4 (94H) analogue.	3.35e-31
36	423	12.5	543	1 R28554	CR1-4 (92T, 94H) analo	4.26e-31
37	422	12.5	543	1 R28559	CR1-4 (114-117STRP, 12	5.41e-31
38	422	12.5	543	1 R28543	CR1-4 (35E, 37Y) analo	5.41e-31
39	416	12.3	543	1 R28546	CR1-4 (44T, 47D, 49L)	2.26e-30
40	411	12.2	579	1 W39924	Amino acid sequence of	7.44e-30
41	390	11.6	830	1 R65216	p-selectin.	1.09e-27
42	389	11.5	577	1 W06882	Membrane co-factor pro	1.38e-27
43	379	11.2	263	1 P92003	Deduced sequence of co	1.48e-26
44	363	10.8	254	1 R47154	Sequence of soluble co	6.46e-25
45	363	10.8	254	1 R47155	Sequence of soluble co	6.46e-25

ALIGNMENTS

RESULT 1
 ID W39154 standard: Protein: 240 AA.
 AC W39154;

DE 27-APR-1998 (first entry)
 DE Human partial Complement factor H protein fragment 1.

KW Complement factor H; tumour associated antigen; renal cancer;
 KM urogenital cancer; medicament; modulator.

OS Homo sapiens.
 PN MO9738136-A1.

PD 16-OCT-1997.
 PF 09-APR-1997: US-812481.

PR 06-MAR-1997: US-812481.
 PR 09-APR-1996: US-015083.

PR 09-APR-1996: US-630048.
 PR 06-MAR-1997: US-038614.

PA (BARD-) BARD DIAGNOSTIC SCI INC.
 PI Enfield DL, Hass GM, Kinders RJ;

DR WPI: 97-512742/47.
 DR N-PSDB: V02790.

PT Treating or screening for cancer, e.g. renal or urogenital cancer -
 PT by modulating or detecting tumour associated human complement factor

PS H related antigen, or nucleic acid encoding it
 PS Example 6B: Fig 6B: 104pp: English.

CC This partial protein sequence represents a region of the human
 CC tumour-associated complement factor H (CFH). This sequence is used

CC in the identification of complement factor H related proteins and
 CC antigens isolated from clone PRB9PH410 (see W39155). The detection of

CC such proteins and a CFH antigens can be used in screening or for the
 CC treatment of renal or urogenital cancer, e.g. bladder, cervical or

CC prostate cancer. Agents that may modulate this antigen could be used in
 CC the manufacture of a medicament for the treatment of a tumour cell.

Sequence 240 AA:
 Query Match 41.1%; Score 1385; DB 1; Length 240;
 Best Local Similarity 71.7%; Pred. No. 1.86e-134;
 Matches 172; Conservative 30; Mismatches 38; Indels 0; Gaps 0;

DB	1	FTLTGNAVFEYGVKAVYVCNGYQLGSLINREDDTGWTNDIPICVYVCLPYTAPENG 60
QY	78	FLIAGSEFEFGAKVYVYCDGTYOLGIDYREDADADMTDIPICEVVKCLPYTELNG 137
DB	61	KIVSAMPEDRREYHFGVAVFYVCNGYKIEGDEDMHCSDDGFMSEKPKCYEISCKSPDV 120
QY	138	KIVSAAEPDQRYTGQVYVREDCNGFKIEGOKEMHSENLMSKPKQCYEISCLPRV 197
DB	121	INGSPISQKIIVKENRFQYKCNMGYERSGDAVCTSGRPLPSCDEKSCNDPIYLPNG 180

QY 198 ENGDIYKPKYKENERFOYKCKGFGYKRGDAVCTGSGMNPQSCHEMTCTFYIPNG 257

Db 181 DYSPLRIKRTGETTETTCORNGFYPATRGNAKCTSTGWIAPRCLTKPCDPYDITKGGGL 240

QY 258 IYTHRIKRIHRIIDDEIRYECCKGFPATRSVSKCTITGWIAPRCLTKPCDPFPQKHGRL 317

RESULT 2

ID W39155 standard; Protein; 216 AA.

AC W39155;

DT 27-Apr-1998 (first entry)

DE Clone PRBB9FH410 CFH related protein fragment.

KM Complement factor H; tumour associated antigen; renal cancer;

KW urogenital cancer; medicament; modulator.

OS Synthetic.

PN MO9738136-A1.

PD 16-OCT-1997.

PF 09-Apr-1997; U05710.

PR 06-MAR-1997; US-812481.

PR 08-APR-1996; US-015083.

PR 09-APR-1996; US-630048.

PR 06-MAR-1997; US-038614.

PA (BARD-) BARD DIAGNOSTIC SCI INC.

PI Enfield DL, Hass GW, Kinders RJ;

DR WPI: 97-512742/47.

DR N-PSDB; V02791.

PT Treating or screening for cancer, e.g. renal or urogenital cancer -

PT by modulating or detecting tumour associated human complement factor

PT H related antigen, or nucleic acid encoding it

PS Example 6B; Fig 6B; 104pp; English.

CC This partial protein is found in clone PRBB9FH410 and represents a

CC complement factor H related protein with homology to a region of the

CC human tumour associated complement factor H (CFH). The detection of this

CC protein and a CFH antigen can be used in screening or for the treatment

CC of renal or urogenital cancer, e.g. bladder, cervical or prostate cancer.

CC Agents that may modulate this antigen could be used in the manufacture of

CC a medicament for the treatment of a tumour cell.

SQ Sequence 216 AA;

Query Match 37.9%; Score 1276; DB 1; Length 216;

Best Local Similarity 73.1%; Pred. No. 1,44e-122;

Matches 158; Conservative 26; Mismatches 32; Indels 0; Gaps 0;

Db 1 TCNEGQQLGEINRYRCDDTGTNDIPICVYKCLPTAPENGRKIVSSANPREHFGQ 60

QY 95 TDEGQQLGEIDIRYRCDDAGTNDIPICVYKCLPTAPENGRIVSGAEPDQYFFQ 154

Db 61 AVRFVNSGKIEGDEHMCSDDGFWGKPKVCVEISCKSPVDINGSPIISOKIYKENER 120

QY 155 VYRFECNSGFKIEGQKEMHCSENGLMSNEKPCQVEISCLPPRENDGDIYKRYKENER 214

Db 121 FOYKCMAGYESRGDAVCTGSGMNPQSCHEMTCTFYIPNGDIYKRYKENER 180

QY 215 FOYKCMAGYESRGDAVCTGSGMNPQSCHEMTCTFYIPNGDIYKRYKENER 274

Db 181 OCRNGFYPATRGNAKCTSTGWIAPRCLTKPCDPY 216

QY 275 ECKNGFYPATRSVSKCTITGWIAPRCLTKPCDPF 310

RESULT 3

ID Y09065 standard; Protein; 578 AA.

AC Y09065;

DT 06-JUL-1999 (first entry)

DE Human complement factor H homolog protein.

KM Human complement factor H; immunological mechanism; complement reaction;

KW gene therapy; immune stimulation; haematopoiesis regulation; chemotactic;

KW tissue growth activity; anti-inflammatory; tumour inhibition;

KW secretory signal.

OS Homo sapiens.

PN MO9918200-A1.

PD 15-APR-1999.

PF 02-OCT-1998; J044448.

PR 06-OCT-1997; JP-272837.

PA (PROT-) PROTEGENE INC.

PA (SAGA-) SAGAMI CHEM RES CENT.

PI Kato S, Sekine S;

DR WPI: 99-264019/22.

DR N-PSDB: X34737.

PT Human proteins with secretory signal sequences and nucleotide

PT sequences, useful in control of proliferation and differentiation of

PT cells

PS Claims 1; Page 55-58; 71pp; English.

CC This DNA encodes a protein having homology to human complement factor H,

CC which plays a role in the immunological mechanism involving the

CC complement reaction. The protein can also be used as an antigen for

CC preparing antibodies against the protein. The cDNA can be used as a probe

CC for gene diagnosis and the gene for gene therapy, as well as for large-

CC scale expression of the protein. The protein may also have immune

CC stimulating or suppressing activity, haematopoiesis regulating activity,

CC tissue growth activity, active/inhibin activity, anti-inflammatory

CC activity, tumour inhibition activity, chemotactic/chemokinetic activity,

CC receptor/ligand activity, etc. The protein is identified by the presence

CC of a hydrophobic N-terminal secretory signal region. Knowledge of the

CC protein function is not required, as in e.g. methods of expression

CC cloning.

SQ Sequence 578 AA;

Query Match 21.1%; Score 712; DB 1; Length 578;

Best Local Similarity 35.1%; Pred. No. 1,36e-61;

Matches 124; Conservative 72; Mismatches 127; Indels 30; Gaps 25;

Db 54 YYCDNFTVPSGSYWDYIHCTODGMSPTVP-C-LRTCSKSDVIEENG-FIS--ESSSY 107

QY 94 YTCDEGQQL-GRD-DYRECDADGWTNDIPICVYKCL-PTVLENGRIVSGAEPDQY 150

Db 108 ILMNEFOYKPKRYAARDNGSSSTICLQNG-WSTQ-PTCIKF-CMPVPENSRAKS-NG 163

QY 151 YFEQVYVRFECNSGF-KIEQKE--MHCSSEGLMSNKPQCVESICLPPVENDGDIYKRP 207

Db 164 MMEKLDITLDYECYDESSYGTNTDSIVCGEDGWSHLPTCVNSSSCGPPPIISNGDTT 223

QY 208 V-YKENERFOYKCKGFG--YKENERGA-VCTGSGMNPQSC-FEM-TC-LTYIPNGIYT 260

Db 224 SFQKYVILPMSRYEYVQCS-YTELQSKYVTCNSGDSPPRCISMKPCEPFIQGHLY 282

QY 261 PHRIKRIHRI-DEIRYECCKGFPATRSVSKCTITGWIAPRCLTKPCDPFPQKHGRL 318

Db 283 YENTRPYPVATGOSYXYCDNFTVPSGSYWDYIHCTODGMSPTVP-C-LRTCSKSDI 342

QY 319 YVESRPYPVPIGKEYSTYICDNGFTTTSQSYWDYLRCTVNGMEVEPCLRQCI-FHY-V 376

Db 343 ENGFTSESSSIYILNKEIYKCKPGYATADGNSGSIITCLQNGMSAOPICIKF 395

QY 377 EYGESSYWMQRRYIEGQSAKVQCHSGYSLPNQGT--YCTENGMSPPPRCVRI 427

RESULT 4

ID R13490 standard; Protein; 581 AA.

AC R13490;

DT 30-OCT-1991 (first entry)

DE Human C4 binding protein.

KW C4bp; monomer; complement protein; p10D.C4bp.3; SCR;

KW short consensus repeat.

OS Homo sapiens.

FT Key

FT peptide

FT 1..32 Location/Qualifiers

FT /label= signal_peptide

FT 33..581

FT /label= c4bp

FT 33..93

FT /label= SCR8

FT 94..155

FT /label= SCR7

FT 156..219

FT /label= SCR6

FT region

FT region

FT region


```

FT region 220..279
FT /label= SCRS
FT region 280..345
FT /label= SCR4
FT region 346..406
FT /label= SCR3
FT region 407..464
FT /label= SCR2
FT region 465..523
FT /label= SCR1
FT domain 524..581
FT /label= C4bp_core
FT /note= "responsible for multimer assembly"
FT disulfide_bond 34..80
FT /note= "intradomain"
FT disulfide_bond 65..92
FT /note= "intradomain"
FT WO9111461-A.
PN 08-AUG-1991.
PF 28-JAN-1991: U00567.
PR 26-JAN-1990: US-470888.
PA (BIOG-) BIOGEN INC.
PI Paszek MP, Winkler G, Liu TR:
DR N-PSDB; Q13242.
PT New C4 binding protein fusion proteins and DNA encoding them -
PT comprise assemblies of C4bp monomers linked to functional moiety,
PT e.g. AbZ, useful as delivery vehicles in diagnosis and therapy
PS Example 1; Fig 1; 105pp; English.
CC This sequence was deduced from human hepatocyte (Hep G2) cDNA
CC obtained following PCR amplification. The protein is a monomer
CC containing 8 SCRs. Each SCR forms a looped domain due to the
CC presence of two intradomain disulphide bonds (only the disulphide
CC bonds of SCR8 are labelled in the Features Table). Within each SCR,
CC the first cysteine residue bonds with the third and the second
CC cysteine residue bonds with the fourth. This secondary structure is
CC responsible for the conformational flexibility of the C4bp monomer.
CC The invention covers fusion proteins in which the monomer sequence,
CC or a specified subfragment of it, e.g. having 5, 4, 3 or 1 SCR(s)
CC is fused to the C-terminal of a protein such as a viral receptor,
CC cell ligand, a bacterial, viral or parasitic immunogen, enzyme,
CC cytokine, toxin, etc. See also Q13243-51.
SQ Sequence 581 AA:

Query Match 15.9%; Score 537; DB 1; Length 581;
Best Local Similarity 30.7%; Pred. No. 5,15e-43;
Matches 138; Conservative 85; Mismatches 182; Indels 45; Gaps 40:

DB 33 NC-GPPTLSFAPAMDITLRETRFKTGLTKYTCLPGLVNSHSTQTLTNSDGEWV-YNT 90
   :| | | | : : : : :| | | | | : : : : :| | | | |
QY 2 DCKGPPPREN-SEILSGSWSEOLYSEGTQAYKCRPGY-RTLGIVYVCK-NGEMVSPNP 58
   :| | | | : : : : :| | | | | : : : : :| | | | |
DB 91 FCILY-KR-CHHPGELRNG--OVEIKTDLSPGSOIEFSGSEFFLTSTSCQVDKRGV 146
   :| | | | | : : : : :| | | | | : : : : :| | | | |
QY 59 SRICKRRPCCGHPGDTGPGSEFLAVGSEPHGAKAVYTCDDESYQLGELDR-EC-DAD-G 115
   :| | | | | : : : : :| | | | | : : : : :| | | | |
DB 147 WSHPLPOCEIYKCKPPDIRNGR-HSG--E-ENFYAYGFSVYSCDPRFSLGHASISCT 202
   :| | | | | : : : : :| | | | | : : : : :| | | | |
QY 116 WTNDIPICEVYKCLPVELENGRIVSGAEPDQGYFGOVVRFCONSGFKIEGKEMHS 175
   :| | | | | : : : : :| | | | | : : : : :| | | | |
DB 203 VENETIGWRSPPTCEKITCRKPDVSHGEMVSGFPIYNAKDTIVRCQKGFVLKRGSSV 262
   :| | | | | : : : : :| | | | | : : : : :| | | | |
QY 176 -EN--GLMSNEKPOCVIEISCLPRVENGDIY-LKPYKENERFOYKCKGFFYKKEG 230
   :| | | | | : : : : :| | | | | : : : : :| | | | |
DB 263 IHCADSKMNPSPACPEINCLPDIPIHASEMYRPTKEDVYVGVLYRCHPGYK 322
   :| | | | | : : : : :| | | | | : : : : :| | | | |
QY 231 AVC-TGSGWNP-QPSCEMTCILT-PIYPNGIY-I-PAIKIRI--DDEI-RIECKNGY 282
   :| | | | | : : : : :| | | | | : : : : :| | | | |
DB 323 TTDEPTIVICOKNLRWTPYOGCEALCCPEKLNNGEIT-QH-RKS-RPANHC-VY-FYGD 377
   :| | | | | : : : : :| | | | | : : : : :| | | | |
QY 283 ATRSPVVS-KCIT-I-GWIPAPRCSILKPCDFQPFKRGRLYESSRRRYPFVIGKEYSYCD 340
   :| | | | | : : : : :| | | | | : : : : :| | | | |
DB 378 EISCHETSFSAL-COGDGTMSPTSCGDICNFP-PKIAHGHIKSSSSSFFKEEII 435

```

```

QY 341 N-GFTTSPQSQSYWDYINCTYNG--WEPEVP-CLRQCLFHYVEYGESSYWR-R-ITEGOSAKV 396
   :| | | | | : : : : :| | | | | : : : : :| | | | |
DB 436 YECDDKGYILV-GQAKLSCYSYSHWSAPAPOC 464
   :| | | | | : : : : :| | | | | : : : : :| | | | |
QY 397 -QCHSGYSLPNQODITYCTENGKSP-PKC 424
   :| | | | | : : : : :| | | | | : : : : :| | | | |

RESULT 5
ID R28557 standard; peptide; 543 AA.
AC R28557:
DT 19-MAR-1993 (first entry)
DE CRI-4 (99H, 103E) analogue.
KW short consensus repeat; regulator of complement activation;
KW C3b binding; C4b binding; human complement type 1 receptor.
OS Homo sapiens.
FH Key Location/Qualifiers
FT region 1..60
FT /label= SCR-1
FT region 61..122
FT /label= SCR-2
FT region 451..510
FT /label= SCR-8
FT region 511..543
FT /label= SCR-9
FT /note= "TRUNCATED"
FT misc-difference 99
FT /note= "Ser substituted by His (SCR-9)"
FT misc-difference 103
FT /note= "Thr substituted by Glu (SCR-9)"
FT EP-512733-A.
PN 11-NOV-1992.
PF 28-APR-1992: 303826.
PR 03-MAY-1991: US-695514.
PA (UNIV ) UNIV WASHINGTON.
PI Atkinson JP, Hourcade D, Kiyoh M;
DR WPI; 92-375009/46.
PT Complement activity regulator protein analogues - useful for
PT treating auto-immune diseases, to suppress transplant rejection,
PT for diagnosis etc.
PS Claim 11; Fig 2 and R11810; 23pp; English.
CC The cDNA clone designated CRI-4 was described in J Exp Med. (1988)
CC 168:1255-1270. It encodes the first 8 and a half amino terminal
CC SCRs of CRI. The invention concerns analogues of "regulator of
CC complement activation" proteins or truncated, hybrid or recombinant
CC forms of them. CRI-4 is a preferred truncated form and a number of
CC specified substitution variants of it are claimed in which certain
CC positions in SCR-2 which have been identified as important for the
CC degree of C3b- and C4b-binding have been substituted by amino acids from
CC the corresponding positions in SCR-9. The specification does not
CC contain the CRI-4 sequence; the sequence given here was constructed
CC from the full-length CRI amino acid sequence having GEMNSEQ
CC accession number R11810 and descriptions in the disclosure.
SQ Sequence 543 AA:

Query Match 13.2%; Score 445; DB 1; Length 543;
Best Local Similarity 32.5%; Pred. No. 2,21e-33;
Matches 105; Conservative 72; Mismatches 106; Indels 40; Gaps 34:

DB 22 PPIGTLYNECRPGYSGR-PSIICLNKSNVWTGAK-DR-CRRKSCRNPDP-PVNGM-VHV 76
   :| | | | | : : : : :| | | | | : : : : :| | | | |
QY 24 YSEGTQATYKCRPRTILGTIVKVC-KNGEMVSPNSRICKRRCGHPGDTPIESSFLAV 82
   :| | | | | : : : : :| | | | | : : : : :| | | | |
DB 77 IKGIQFSQIKYSCTKGYRLIGHSS-AECIISGDTIWMNETPICDRIPGLPPT-ITNG 134
   :| | | | | : : : : :| | | | | : : : : :| | | | |
QY 83 GSEPEFGAKVYVYTCDEGQILGEIDYREC--DADG--WTNDIPICEVYKCLPVTLENG 137
   :| | | | | : : : : :| | | | | : : : : :| | | | |
DB 135 DFIS--INRE-NHIYGSVYTRCNPGSGGRVFLVPEPSTICYNSNDQVQIMSGPAPOC 191
   :| | | | | : : : : :| | | | | : : : : :| | | | |
QY 138 RIVSGAEPDDEYFGOVVRECN-SGF-KI-E-GQKEHNC-S-EN-GLMSNEKPOC 187
   :| | | | | : : : : :| | | | | : : : : :| | | | |
DB 192 IIPNKCTPVENENGLVSDNRSLSLNEVVEFRQCPVVMGPRRVYCOALNKEPELPS 251
   :| | | | | : : : : :| | | | | : : : : :| | | | |

```

QY 188 VEIS-CLPVRVNGDGIY-LKPYKENERFOYKCKOGFYKBERGAVCTG-SGMNPQ-PS 243
DB 252 CSRV-CQPPDVLHAERORDKNSPGGEVYVYSCPEP-Y-DLRGAASRCRCPQGDMSFA 308
QY 244 CEMMCLTP-YIPNGITYPH-RIKHRIDELINECKNGFYPATRSFVS-KCTITG-WIPA 299
DB 309 APTCEVKSQDPMGQLNGRVLE 331
QY 300 -PRCSLKPCD-F-PQFKHGRLYX 319

RESULT 6

ID R28552 standard; peptide; 543 AA.
AC R28552;
DT 19-MAR-1993 (first entry)
DE CRI-4 (78T, 79D) analogue.
KW short consensus repeat; regulator of complement activation;
KW C3b binding; C4b binding; human complement type 1 receptor.
OS Homo sapiens.

FT Key Location/Qualifiers
FT region 1..60
FT /label=SCR-1
FT region 61..122
FT /label=SCR-2
FT region 451..510
FT /label=SCR-8
FT region 511..543
FT /label=SCR-9
FT misc_difference 78
FT /note= "TRUNCATED"

FT misc_difference 79
FT /note= "Lys substituted by Thr (SCR-9)"
FT misc_difference 79
FT /note= "Gly substituted by Asp (SCR-9)"
PN EP-512733-A.
PD 11-NOV-1992.
PE 28-APR-1992; 303826.
PR 03-MAY-1991; US-695514.
PA (UNIM) UNIV WASHINGTON.
PI Atkinson JP, Hourcade D, Knych M;
DR WPI: 92-375009/46.

PT Complement activity regulator protein analogues - useful for
PT treating auto-immune diseases, to suppress transplant rejection,
PT for diagnosis etc.
PS Claim 11: Fig 2 and R11810: 23pp; English.
CC The cDNA clone designated CRI-4 was described in J.Exp.Med.(1988)
CC 168:1255-1270. It encodes the first 8 and a half amino terminal
CC SCR's of CRI. The invention concerns analogues of "regulator of
CC complement activation" proteins or truncated, hybrid or recombinant
CC forms of them. CRI-4 is a preferred truncated form and a number of
CC specified substitution variants of it are claimed in which certain
CC positions in SCR-2 which have been identified as important for the
CC degree of C3b- and C4b-binding are substituted by amino acids from
CC the corresponding positions in SCR-9. The specification does not
CC contain the CRI-4 sequence; the sequence given here was constructed
CC from the full-length CRI amino acid sequence having GENSEQ
CC accession number R11810 and descriptions in the disclosure.
SQ Sequence 543 AA;

Query Match 13.2%; Score 445; DB 1; Length 543;
Best Local Similarity 32.0%; Pred. No. 2,21e-33;
Matches 103; Conservative 74; Mismatches 107; Indels 38; Gaps 33;

DB 22 FPIGTLYNTEGRPGYSGR-PFSITCLKNSVWTGAK-DR-CRRKSCRNPDP-PVNGM-VHV 76
QY 24 YSECTQATYTCRCRGYRLGTYVVC-KNGEWPVSPSRICRRCPCGPGTTPGSFRLAV 82
DB 77 ITDIQFSQIKYSCRTKGRIGSSATCIISGDTVIWNETPDCRIPCGLPPT-ITNGD 135
QY 83 GSFEFGAKVYVYTCDBGYQLGEIDYR-ECDDAG--WTNDIPICEVVC-LPTYLENGR 138
DB 136 FIS--TNRE-NFHGVSYYTRCNPBGSGKRVFELVGPSTYCTSNDDQVIGSGAPQCI 192
QY 139 IVSGAALPDDQYFGGVYREC--SGF-KI-E--GQKEMHC-S-EN--GLWMSNEKPCV 188

DB 193 IPNKTPPNVEGLIVSDNLSFLINEVFERCQPVFMKGPFRVYKCOALMKWELPSC 252
QY 189 EIS-CLPVRVNGDGIY-LKPYKENERFOYKCKOGFYKBERGAVCTG-SGMNPQ-PSC 244
DB 253 SRV-CQPPDVLHAERORDKNSPGGEVYVYSCPEP-Y-DLRGAASRCRCPQGDMSFA 309
QY 245 EEMTCLTP-YIPNGITYPH-RIKHRIDELINECKNGFYPATRSFVS-KCTITG-WIPA 299
DB 310 PTCEVKSQDPMGQLNGRVLE 331
QY 300 -PRCSLKPCD-F-PQFKHGRLYX 319

RESULT 7

ID W73147 standard; protein; 778 AA.
AC W73147;
DT 29-JAN-1999 (first entry)
DE Amino acid sequence of the soluble complement receptor 1 (SCR1).
KW Human: soluble complement receptor 1; SCR1; T-cell; B-cell.
KW mediated immune response; inhibition; tissue rejection; gene therapy;
KW dystrophin; inflammatory response; interferon-gamma secretory response;
KW autoimmune response; neurological response; Alzheimer's disease;
KW Parkinson's disease; multiple sclerosis; systemic lupus erythematosus;
KW rheumatoid arthritis; myasthenia gravis; epidermis bullosa;
KW Hashimoto's disease.
OS Homo sapiens.
PN W09845430-A1.
PD 15-OCT-1998.
PE 06-APR-1998; G01012.
PR 05-APR-1997; GB-006950.
PA (ANNE/) ANNENKOV A.
PI (CHER/) CHERNAJOVSKY Y.

DR WPI: 98-568350/48.
PT Fragment of soluble human complement receptor 1 - useful for
PT treating T-cell or B-cell mediated immune responses e.g.
PT inflammatory responses such as rheumatoid arthritis
PS Disclosure; Fig 1; 54pp; English.
CC This is an amino acid sequence of the human soluble complement
CC receptor 1 (SCR1), useful in the treatment of T-cell or B-cell
CC mediated immune responses. It is used to inhibit a T-cell or
CC B-cell-mediated immune response to prevent immune response-mediated
CC tissue rejection and destruction or clearance or inactivation of an
CC expressed protein especially from cells that have been treated by gene
CC therapy to express the protein, e.g. dystrophin. The protein can also
CC be used to inhibit a T-cell or B-cell-mediated inflammatory response,
CC an interferon-gamma secretory response, autoimmune response or
CC neurological response, e.g. Alzheimer's or Parkinson's disease or
CC multiple sclerosis. Also the protein can be used to treat systemic
CC lupus erythematosus, rheumatoid arthritis, myasthenia gravis,
CC epidermis bullosa or Hashimoto's disease.
SQ Sequence 778 AA;

Query Match 13.2%; Score 446; DB 1; Length 778;
Best Local Similarity 32.3%; Pred. No. 1.74e-33;
Matches 104; Conservative 72; Mismatches 108; Indels 38; Gaps 33;

DB 68 FPIGTLYNTEGRPGYSGR-PFSITCLKNSVWTGAK-DR-CRRKSCRNPDP-PVNGM-VHV 122
QY 24 YSECTQATYTCRCRGYRLGTYVVC-KNGEWPVSPSRICRRCPCGPGTTPGSFRLAV 82
DB 123 IKGIQFSQIKYSCRTKGRIGSSATCIISGDTVIWNETPDCRIPCGLPPT-ITNGD 181
QY 83 GSFEFGAKVYVYTCDBGYQLGEIDYR-ECDDAG--WTNDIPICEVVC-LPTYLENGR 138
DB 182 FIS--TNRE-NFHGVSYYTRCNPBGSGKRVFELVGPSTYCTSNDDQVIGSGAPQCI 238
QY 139 IVSGAALPDDQYFGGVYREC--SGF-KI-E--GQKEMHC-S-EN--GLWMSNEKPCV 188
DB 239 IPNKTTPNVEGLIVSDNLSFLINEVFERCQPVFMKGPFRVYKCOALMKWELPSC 298
QY 189 EIS-CLPVRVNGDGIY-LKPYKENERFOYKCKOGFYKBERGAVCTG-SGMNPQ-PSC 244

AC R28553;
 DT 19-MAR-1993 (first entry)
 DE CRI-4 (85R, 87N) analogue.
 KW short consensus repeat; regulator of complement activation;
 OS Homo sapiens.
 FH C3b binding; C4b binding; human complement type 1 receptor.
 FT Key
 FT region Location/Qualifiers
 FT 1..60
 FT /label= SCR-1
 FT 61..122
 FT /label= SCR-2
 FT 451..510
 FT /label= SCR-8
 FT 511..543
 FT /label= SCR-9
 FT /note= "TRUNCATED"
 FT misc_difference 85
 FT /note= "Gln substituted by Arg (SCR-9)"
 FT misc_difference 87
 FT /note= "Lys substituted by Asn (SCR-9)"
 PD EP-512733-A.
 PD 11-NOV-1992.
 PR 28-APR-1992; 303826.
 PR 03-MAY-1991; US-695514.
 PA (UNIW) UNIV WASHINGTON.
 PI Atkinson JP, Hourcade D, Krych M;
 DR WPI: 92-375009/46.
 PT Complement activity regulator protein analogues - useful for
 PT treating auto-immune diseases, to suppress transplant rejection,
 PT for diagnosis etc.
 PS Claim 11; Fig 2 and R11810; 23pp; English.
 CC The cDNA clone designated CRI-4 was described in J.Exp.Med.(1988)
 CC 168:1255-1270. It encodes the first 8 and a half amino terminal
 CC SCRs of CRI. The invention concerns analogues of "regulator of
 CC complement activation" proteins or truncated, hybrid or recombinant
 CC forms of them. CRI-4 is a preferred truncated form and a number of
 CC specified substitution variants of it are claimed in which certain
 CC positions in SCR-2 which have been identified as important for the
 CC degree of C3b- and C4b-binding are substituted by amino acids from
 CC the corresponding positions in SCR-9. The specification does not
 CC contain the CRI-4 sequence; the sequence given here was constructed
 CC from the full-length CRI amino acid sequence having GENSESO
 CC accession number R11810 and descriptions in the disclosure.
 SQ Sequence 543 AA;

Query Match 13.1%; Score 441; DB 1; Length 543;
 Best Local Similarity 32.0%; Pred. No. 5,77e-33;

Matches 103; Conservative 73; Mismatches 108; Indels 38; Gaps 33;

DB 22 PPIGTLYNTECRPGYSGR-PFSITCLKNSVWTGAK-DR-CRRKSCNPPD-PVNGM-VHY 76
 QY 24 YSEGTATYKCRPGYRFLGTIVKVC-KNGEWPVNSNRICKRKPCHPDPTFGSRLAV 82
 DB 77 IKGIQFGRINYSCTKGYRLIGSSSATCIISGPTVIMNDETPICDRIPGCLPPT-ITNGD 135
 QY 83 GSEFEFGAKVYTCDEGYQLGEIDYR-ECDAAG--WTNDIPCEVVKC-LPVTELENGR 138
 DB 136 FIS--TNRE-NHYGSVYTYRCNPGSGRKRVFELVGEPSITCTSNDDQVIGISGAPOCI 192
 QY 139 IVSGAEPDQYEFYFGVYFEFCN--SGF-KI-E--GQKMH-C-S-EN--GLMSNKPQCV 188
 DB 193 IPKRCPPVNEGILVSDNRSLSFSLNEVEVEFRQPYVVKGPFRVYKCOALNKEPELPSC 252
 QY 189 EIS-CLPPVNEVDGTY-LKPYKENERFOYKCKQGFYKKEGDAVCG-SCWNQ-PSC 244
 DB 253 SRV-CQPPDVLAERTQDKDNFSPGOEVFYSCEPG-Y-DLGAASMRCTPDGWSPPAA 309
 QY 245 EEMTCLTP-YIINGITTPH-RIKHRIDDIRCKNGEFPATRSVPS-KCTITG-WIPA- 299
 DB 310 PTCEVKSQCDDEMGQLNGRVLF 331
 QY 300 PRCSLKPCD-F-PQFKHGRLY 319

RESULT 11
 ID R28560 standard; peptide; 543 AA.
 AC R28560;
 DT 19-MAR-1993 (first entry)
 DE CRI-4 (114S) analogue.
 KW short consensus repeat; regulator of complement activation;
 OS Homo sapiens.
 FH C3b binding; C4b binding; human complement type 1 receptor.
 FT Key
 FT region Location/Qualifiers
 FT 1..60
 FT /label= SCR-1
 FT 61..122
 FT /label= SCR-2
 FT 451..510
 FT /label= SCR-8
 FT 511..543
 FT /label= SCR-9
 FT /note= "TRUNCATED"
 FT misc_difference 114
 FT /note= "Asp substituted by Ser (SCR-9)"
 PD EP-512733-A.
 PD 11-NOV-1992.
 PR 28-APR-1992; 303826.
 PR 03-MAY-1991; US-695514.
 PA (UNIW) UNIV WASHINGTON.
 PI Atkinson JP, Hourcade D, Krych M;
 DR WPI: 92-375009/46.
 PT Complement activity regulator protein analogues - useful for
 PT treating auto-immune diseases, to suppress transplant rejection,
 PT for diagnosis etc.
 PS Claim 11; Fig 2 and R11810; 23pp; English.
 CC The cDNA clone designated CRI-4 was described in J.Exp.Med.(1988)
 CC 168:1255-1270. It encodes the first 8 and a half amino terminal
 CC SCRs of CRI. The invention concerns analogues of "regulator of
 CC complement activation" proteins or truncated, hybrid or recombinant
 CC forms of them. CRI-4 is a preferred truncated form and a number of
 CC specified substitution variants of it are claimed in which certain
 CC positions in SCR-2 which have been identified as important for the
 CC degree of C3b- and C4b-binding are substituted by amino acids from
 CC the corresponding positions in SCR-9. The specification does not
 CC contain the CRI-4 sequence; the sequence given here was constructed
 CC from the full-length CRI amino acid sequence having GENSESO
 CC accession number R11810 and descriptions in the disclosure.
 SQ Sequence 543 AA;

Query Match 13.0%; Score 439; DB 1; Length 543;
 Best Local Similarity 32.0%; Pred. No. 9,31e-33;

Matches 103; Conservative 73; Mismatches 108; Indels 38; Gaps 33;

DB 22 PPIGTLYNTECRPGYSGR-PFSITCLKNSVWTGAK-DR-CRRKSCNPPD-PVNGM-VHY 76
 QY 24 YSEGTATYKCRPGYRFLGTIVKVC-KNGEWPVNSNRICKRKPCHPDPTFGSRLAV 82
 DB 77 IKGIQFGRINYSCTKGYRLIGSSSATCIISGPTVIMNDETPICDRIPGCLPPT-ITNGD 135
 QY 83 GSEFEFGAKVYTCDEGYQLGEIDYR-ECDAAG--WTNDIPCEVVKC-LPVTELENGR 138
 DB 136 FIS--TNRE-NHYGSVYTYRCNPGSGRKRVFELVGEPSITCTSNDDQVIGISGAPOCI 192
 QY 139 IVSGAEPDQYEFYFGVYFEFCN--SGF-KI-E--GQKMH-C-S-EN--GLMSNKPQCV 188
 DB 193 IPKRCPPVNEGILVSDNRSLSFSLNEVEVEFRQPYVVKGPFRVYKCOALNKEPELPSC 252
 QY 189 EIS-CLPPVNEVDGTY-LKPYKENERFOYKCKQGFYKKEGDAVCG-SCWNQ-PSC 244
 DB 253 SRV-CQPPDVLAERTQDKDNFSPGOEVFYSCEPG-Y-DLGAASMRCTPDGWSPPAA 309
 QY 245 EEMTCLTP-YIINGITTPH-RIKHRIDDIRCKNGEFPATRSVPS-KCTITG-WIPA- 299
 DB 310 PTCEVKSQCDDEMGQLNGRVLF 331
 QY 300 PRCSLKPCD-F-PQFKHGRLY 319

ID	Accession	Standard	Peptide	AA
RESULT	12			
AC	R28558	standard:	peptide:	543 AA.
DT	19-MAR-1993	(first entry)		
DE	CRI-4 (109N, 110A, 111A, 112H) analogue.			
KW	short consensus repeat; regulator of complement activation; C3b binding; C4b binding; human complement type 1 receptor.			
OS	Homo sapiens.			
FT	Key	Location/Qualifiers		
FT	region	1..60		
FT	region	/label= SCR-1		
FT	region	61..122		
FT	region	/label= SCR-2		
FT	region	451..510		
FT	region	/label= SCR-8		
FT	region	511..543		
FT	region	/label= SCR-9		
FT	misc_difference	/note="TRUNCATED"		
FT	misc_difference	109		
FT	misc_difference	/note="Asp substituted by Asn (SCR-9)"		
FT	misc_difference	110		
FT	misc_difference	/note="Thr substituted by Ala (SCR-9)"		
FT	misc_difference	111		
FT	misc_difference	/note="Val substituted by Ala (SCR-9)"		
FT	misc_difference	112		
FT	misc_difference	/note="Ile substituted by His (SCR-9)"		
PN	EP-512733-A.			
PD	11-NOV-1992.			
PF	28-APR-1992.	303826.		
PR	03-MAY-1991.	US-695514.		
PA	(UNIW) UNIV WASHINGTON.			
PI	Atkinson JP, Hourcade D, Krych W:			
PT	WPI. 92-375009/46.			
PT	Complement activator/regulator protein analogues - useful for treating auto-immune diseases, to suppress transplant rejection, for diagnosis etc.			
PT	for diagnosis etc.			
PS	Claim 11: Fig 2 and R1810: 23pp: English.			
CC	The cDNA clone designated CRI-4 was described in J. Exp. Med. (1988)			
CC	168:11255-1270. It encodes the first 8 and a half amino terminal			
CC	SCRs of CRI. The invention concerns analogues of "regulator of complement activation" proteins or truncated, hybrid or recombinant			
CC	forms of them. CRI-4 is a preferred truncated form and a number of			
CC	specified substitution variants of it are claimed in which certain			
CC	positions in SCR-2 which have been identified as important for the			
CC	degree of C3b- and C4b-binding are substituted by amino acids from			
CC	the corresponding positions in SCR-9. The specification does not			
CC	contain the CRI-4 sequence; the sequence given here was constructed			
CC	from the full-length CRI amino acid sequence having GENSEQ			
CC	accession number R1810 and descriptions in the disclosure.			
SO	Sequence 543 AA:			
Query Match	13.0%;	Score 439;	DB 1;	Length 543;
Best Local Similarity	32.0%;	Pred. No. 9, 31e-33.		
Matches 103;	Conservative	72;	Mismatches 109;	Indels 38; Gaps 33
D2	22	FPKTYLNTYRCRGYSGR--PFSIIICLKNVMTGAK-DR-CRRKSCRPDP--PVNGM-VHV	76	
OY	24	YSEETQATYTCRGRCYRTLGTIVKVC-KNEWMPSPNSRRCRRKPCGHGDPFGFSRLAV	82	
D2	77	IKGIQFGSQIKYSCYKGYRLIGSSSATCISGNAAHMNETYICRIQICGLPPT-ITNGD	135	
OY	83	GSEFEFGAKVYVYFGGVAFECN--IDYREC-DADMTWIDIDICEVWC-LPVELENGR	138	
D2	136	FIS--TKRE-NFHYGSAVYTRCRNPGSGKRVFELVGEPSITYCSDNDQDGIWSPAPCI	192	
OY	139	IVSGAAPPDEYVFGGVAFECN--SGF-KT-E--GQEMHC-S-EN--GLMSNKKPCV	188	
D2	193	IPNKCETPPNENGLIVSDNKSFLNEVVEFRCPYFVYKGGPRRYQALNFWEDLPD	252	
OY	189	EIS-CLPPRYENGGCI-LKPYVKKENREFYKCKGQFVYKKERDAVCTG--SGMNPQ-PSC	244	

D	b	253	SRY-CQPPPVVLAETQTQDXDNFSGDVEFCSPG-Y-DLRGAASRCRPGDSFPAA	309
O	y	245	EEMTCLLP-YIPGVITPH-RIKHRIDDEIRYCKNGEFTPATREVS-KCIIIG-WIPA-	299
D	b	310	PTECVKSCDDFMGLNGRYLE	331
O	y	300	PRCSLKPCD-F-PQFKHGRLYX	319
RESULT 13				
ID	R28571;	standard; peptide; 543 AA.		
AC	R28571;	(first entry)		
DT	19-MAR-1993			
DE	CRI-4 (364-367 NAAH) analogue.			
KW	short consensus repeat; regulator of complement activation;			
KM	C3b binding; C4b binding; human complement type 1 receptor.			
OS	Homo sapiens.			
FH	Key	Location/Qualifiers		
FT	region	1..60		
FT	/label= SCR-1			
FT	region	61..122		
FT	/label= SCR-2			
FT	region	451..510		
FT	/label= SCR-8			
FT	region	511..543		
FT	/label= SCR-9			
FT	/note= "TRUNCATED"			
FT	misc_difference	364..367		
FT	/note= "MSL substituted with NAAH from			
FT	SCR-8-9 to increase C3b binding"			
PN	EP-512733-A.			
PD	11-NOV-1992.			
PF	28-APR-1992; 303826.			
PR	03-MAY-1991; US-695514.			
PA	(UNITM) UNITV WASHINGTON.			
PI	Atkinson JP, Hourcade D, Krych M:			
DR	WPI: 92-375009/46.			
PT	Complement activitiy regulator protein analogues - useful for			
PT	treating auto-immune diseases; to suppress transplant rejection,			
PT	for diagnosis etc.			
PS	Example 8; Page 18 and R11810; 23pp; English.			
CC	The cDNA clone designated CRI-4 was described in J Exp Med.(1988)			
CC	168:1255-1270. It encodes the first 8 and a half amino terminal			
CC	SCRs of CRI. The invention concerns analogues of "regulator of			
CC	complement activation" proteins or truncated, hybrid or recombinant			
CC	forms of them. CRI-4 is a preferred truncated form and a number of			
CC	specified substitution variants of it are disclosed in which certain			
CC	positions in SCR-5-6 are substituted by amino acids from			
CC	the corresponding positions in SCRs which are involved in C3b- and			
CC	C4b-binding. The substitution variant given here has increased C3b-			
CC	binding. The specification does not contain the CRI-4 sequence;			
CC	the sequence given here was constructed from the full-length CRI			
CC	amino acid sequence having GENESED accession number R11810 and			
CC	descriptions in the disclosure.			
SC	Sequence 543 AA:			
	Query Match	13.0%; Score 438; DB 1; Length 543;		
	Best Local Similarity 31.9%; Pred. No. 1,18e-32;			
	Matches 119; Conservative 79; Mismatches 127; Indels 48; Gaps 39;			
D	b	22	FFIISTLVNYECRDPSYGSR-PFSIICLKNSVTWGAK-DR-CRRKSCRNPDP-PVNGM-VHV	76
O	y	24	YSBETOATFYCRCPRGRTLGIVAVC-KNGEWVSPNPSRICRKRPCHGDIDPFGSFRILV	82
D	b	77	IKGIQSFSSQIKYSCSTGYRLIGSSSAATCIISDGYIMWDENPICRIFCGAPPT-ITNGD	135
O	y	83	GSEFEFAKVAVTYTCDESYQLGEIDYR-ECDADG-WINDPIDCEYKC-LPYTELENGR	138
D	b	136	FIS--TVRE-NFHGASVATRCNDPSGGKRVELVGPESITYCTSNDDOYGVMGSPAPOCI	192
O	y	139	IYSGAALPEDEIFYTGQVVIRECN-SGF-KI-E-GQKEMHG-S-EN-GIMSNEKPCV	188
D	b	193	IPNKCTPPNVENGILLVDNSLSLAEVVEFRQPVPFYVAKGRRVRKQCALKMKWEDELPSG	252

QY 189 EIS-CLPFRVENGDDIY-LKPVYKENERFOYCKOGFYKRGDAVCTG-SGMNPQ-PSC 244
 DB 253 SRV-COPPPDVLHAERTORDKDNFSPGOEYFYSCPEG-Y-DLRGAASMRCTPGQDMSPPA 309
 QY 245 EEMTCLTP-YIPNGIYTPH-RIKHRIDELRYECKNGFYPATRSPVS-KCTITG-WIPA- 299
 DB 310 PTCEYKSCD--DEM-GOLL--NGR-VLEPVNLQGLAKVDEVCDEGFOITKGSASV-CYLA 362
 QY 300 PRCSLKPCDDFOFKHGRLYEESRPFYFV--PIGKRYSYCCDNGF--TTPSQSYMDYLR 355
 DB 363 GNAHMSVSPVC 375
 QY 356 CTYNGMEPEVP-C 367

RESULT 14
 ID R28550 standard; peptide; 543 AA.
 AC R28550;
 DT 19-MAR-1993 (first entry)
 DE CRI-4 (64K) analogue.
 KW short consensus repeat; regulator of complement activation;
 KW C3b binding; C4b binding; human complement type 1 receptor.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT region 1..60
 FT region /label=SCR-1
 FT region 61..122
 FT region /label=SCR-2
 FT region 451..510
 FT region /label=SCR-8
 FT region 511..543
 FT region /label=SCR-9
 FT region /note="TRUNCATED"
 FT misc_difference 64 /note="Arg substituted by Lys (SCR-9)"
 PN EP-512733-A.
 PD 11-NOV-1992.
 PF 28-APR-1992; 303826.
 PR 03-MAY-1991; US-695514.
 PA (UNITW) UNIV WASHINGTON.
 PI Atkinson JP, Hourcade D, Krych M;
 DI WPT; 92-375009/46.
 PT Complement activity regulator protein analogues - useful for
 PT treating autoimmune diseases, to suppress transplant rejection,
 PT for diagnosis etc.
 PS Claim 11: Fig 2 and R11810; 23pp; English.
 CC The cDNA clone designated CRI-4 was described in J.Exp.Med.(1988)
 CC 168:1255-1270. It encodes the first 8 and a half amino terminal
 CC SCRs of CRI. The invention concerns analogues of "regulator of
 CC complement activation" proteins or truncated, hybrid or recombinant
 CC forms of them. CRI-4 is a preferred truncated form and a number of
 CC specified substitution variants of it are claimed in which certain
 CC positions in SCR-2 which have been identified as important for the
 CC degree of C3b- and C4b-binding are substituted by amino acids from
 CC the corresponding positions in SCR-9. The specification does not
 CC contain the CRI-4 sequence; the sequence given here was constructed
 CC from the full-length CRI amino acid sequence having GENSSEQ
 CC accession number R11810 and descriptions in the disclosure.
 SQ Sequence 543 AA;

Query Match 13.0%; Score 437; DB 1; Length 543;
 Best Local Similarity 32.0%; Pred. No. 1.50e-32;
 Matches 103; Conservative 72; Mismatches 109; Indels 38; Gaps 33;

DB 22 FPIGTINTECRPGYSGR-PSITICLNKSNVTGAK-DR-CRRKSKRPPD-PVNGM-VHV 76
 QY 24 YSEGTQATYKRCRPRYRLGLTIVKVC-KNGEWPSPNPSIRCKRRCGHPGTGPGSFRILAV 82
 DB 77 IKGIOFSSQIKYSCTKGRILIGSSSATCIISGPTVINDNEPTICDRIPGGLPPT-ITND 135
 QY 83 GSEFEFAKVVYTCDEGTYQLIGEDYR-ECDADG--WTNDIPICEVYKC-LPVTLENGR 138

DB 136 FIS--TNRE-NHYGSVYTYRCPNPGSGGRKYFELVGEPSIYCTSDNDQYIGSGAPOCI 192
 QY 139 IYSGAAEPDQYFYFGQVYFECN--SGF-KI-E--GQEMHC-S-EN--GLMSNEKQCV 188
 DB 193 INKCTPRVENVGLIVSNRSLFSLNEVEYRCQPVFMKGPARRVCOALNKMBELPSC 252
 QY 189 EIS-CLPFRVENGDDIY-LKPVYKENERFOYCKOGFYKRGDAVCTG-SGMNPQ-PSC 244
 DB 253 SRV-COPPPDVLHAERTORDKDNFSPGOEYFYSCPEG-Y-DLRGAASMRCTPGQDMSPPA 309
 QY 245 EEMTCLTP-YIPNGIYTPH-RIKHRIDELRYECKNGFYPATRSPVS-KCTITG-WIPA- 299
 DB 310 PTCEYKSCDDFMGOLLNGRVLE 331
 QY 300 PRCSLKPCD-F-PQKHGRLY 319

RESULT 15
 ID R11982 standard; Protein; 1537 AA.
 AC R11982;
 DT 25-JUN-1991 (first entry)
 DE Partial human complement type 1 receptor.
 KW complement system; C3b/C4b receptor; CRI; allergic reaction;
 KW immune response; long homologous repeat; LHR.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT region 1..438
 FT region /label=LHR-B
 FT region 439..891
 FT region /label=LHR-C
 FT region 892..1341
 FT region /label=LHR-D
 FT region 1495..1498
 FT region /note="positively-charged; preceded by hydrophobic
 FT sequence"
 FT 1521..1526
 FT /note="has 67 per cent homology to site of protein
 FT kinase C phosphorylation in the EGF
 FT receptor"

PN W09105047-A.
 PD 18-APR-1991.
 PF 25-SEP-1990; U05454.
 PR 26-SEP-1989; US-412745.
 PR 26-SEP-1990; US-912349.
 PA (TCEL-) T CEL SCI INC.
 PA (UYJO-) JOHNS HOPKINS UNIVERSITY.
 PA (BRIG-) BRIGHAM AND WOMEN'S HOSPITAL.
 PI Fearon DT, Klickstein LB, Wong WW, Carson GR, Hoh M, Concino MF,
 PI Makrides SC, Marsh HC;
 DI WPT; 91-132854/18.
 DR N-PSDB; Q11643.
 PT Human complement receptor type 1 gene, encoded proteins and
 PT fragments - for treatment of immune disorders, myocardial infarct,
 PT damage due to inflammation and in treatment of thrombosis
 PS Disclosure; Fig 5; 234pp; English.
 CC This sequence comprises three of the four tandem, direct, long
 CC homologous repeats of the full-length F allozyme of CRI. LHR-A is
 CC absent. Each LHR might represent a single C3b/C4b binding domain,
 CC making the receptor multivalent. The LHR's are composed of 7 short
 CC consensus repeats of 60-70 residues resembling the SCR's of other
 CC C3/C4 binding proteins. The protein and fragments of it having C3b
 CC and/or C4b binding activity can be used to treat immune disorders
 CC or disorders involving inappropriate complement activity.
 CC See also Q11642.
 SQ Sequence 1537 AA;

Query Match 13.0%; Score 439; DB 1; Length 1537;
 Best Local Similarity 32.6%; Pred. No. 9.31e-33;
 Matches 105; Conservative 73; Mismatches 106; Indels 38; Gaps 33;

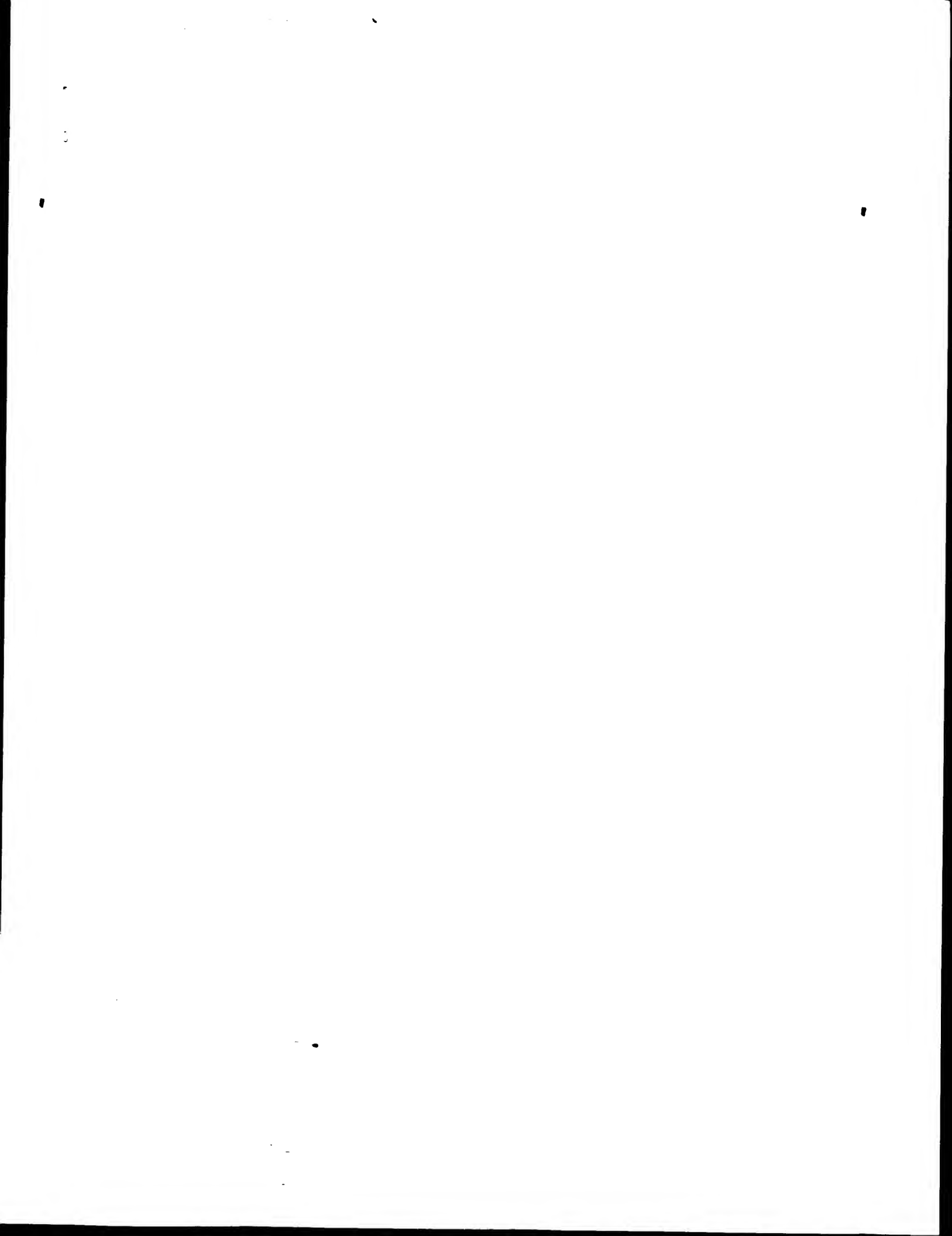
DB 461 FPIGTINTECRPGYSGR-PSITICLD-NLYWSSPKDYCKRCKSKTPPD-PVNGM-VHV 516
 QY 24 YSEGTQATYKRCRPRYRLGLTIVKVC-KNGEWPSPNPSIRCKRRCGHPGTGPGSFRILAV 83

```

Db 517 TDIOVGRINSCITGHRIGHSS-ABCIISGNTAHNSTKRPICQRIPOGLPPT-IANGD 574
QY 84 SEFEFGAKVYTCDEGYQLLGEIDYREC--DAD-G-WTNDIPICEVVKC-LPVTLENGR 138
Db 575 FIS--TNRE-NFHYGSVYTYRCNLGSRGRKRYFELVGEPSIYCTSNDDQVGIWGPAPQCI 631
QY 139 IYSGAABPDQEVYFGQVYRFECH--S-GFKI-E--GQKEMHC-S-EN--GLMSNEKPCV 188
Db 632 IPNKCTPPNVENGIIIVSDNLSLFSINEVEFRPCOPGFYMGPRRYKCOALNKMEBELPSC 691
QY 189 EIS-CLPPREVENGDIX-LKPYKENERFOYKCKOGFYKKEKGDVCTG--SGWNPQ--PSC 244
Db 692 SRV--QPPPEILHGHHTPSHONFSPGOEVFYSCPEG-Y-DLRGAASLHCTPRGDMSPEA 748
QY 245 EEMTCLTP-YIPNGIYTP-HRIKRIIDELIRYCKNGFYPATRSPVS-KCITITG-WIP-A 299
Db 749 PRCAVKSDDDFLGOLPHGRVLF 770
QY 300 PRCSLKPCD-F-PQFKHGRLY 319

```

Search completed: Thu Jun 8 21:49:27 2000
 Job time : 17 secs.



Distribution rights by Oxford Molecular Ltd

Tabular output not generated.

Sequence:

Searched:

Database:

Statistics:

SUMMARIES

23 434

45 345

ALIGNMENTS

DATE _____

ACCESSIONS

REFERENCE

#journal

1
 2
 3
 4
 5
 6

#Cross-Tei
#accesior

##mol_e

##CROSS

REFERENCE

#journal

SECRET

#Cross-Te:

```
##stat
```

```
##resi
```

#CLOS
DEPENDENCE

#authors

#title

#Cross-Te

```
##stat
```

```
##resj
##atom#
```

##CROS

COMMENT F

0
0
:
9
8
6
4
2

```

#map position 1
CLASSIFICATION #superfamily complement factor H; complement factor H repeat
KEYWORDS homology
complement alternate pathway; duplication; glycoprotein;
plasma
FEATURE
1-18 #domain signal sequence #status predicted #label SIG\
19-134 #product complement factor H #status predicted #label
MPV\
21-80 #domain complement factor H repeat homology #label FH01\
85-141 #domain complement factor H repeat homology #label FH02\
146-205 #domain complement factor H repeat homology #label FH03\
210-262 #domain complement factor H repeat homology #label FH03\
246-248 #region cell attachment (R-G-D) motif\
267-320 #domain complement factor H repeat homology #label FH05\
325-385 #domain complement factor H repeat homology #label FH06\
389-442 #domain complement factor H repeat homology #label FH07\
448-505 #domain complement factor H repeat homology #label FH08\
509-564 #domain complement factor H repeat homology #label FH09\
569-622 #domain complement factor H repeat homology #label FH10\
629-683 #domain complement factor H repeat homology #label FH11\
690-743 #domain complement factor H repeat homology #label FH12\
752-802 #domain complement factor H repeat homology #label FH13\
808-861 #domain complement factor H repeat homology #label FH14\
867-931 #domain complement factor H repeat homology #label FH15\
936-989 #domain complement factor H repeat homology #label FH16\
994-1048 #domain complement factor H repeat homology #label FH17\
1053-1107 #domain complement factor H repeat homology #label FH18\
1114-1168 #domain complement factor H repeat homology #label FH19\
1172-1233 #domain complement factor H repeat homology #label FH20\
21-66,52-80,85-129,
114-141,146-192,
178-205,210-251,
237-262,267-309,
294-320,325-374,
357-385,389-431,
416-442,448-494,
477-505,509-553,
536-564,569-610,
597-622,629-672,
658-683,680-732,
718-743,752-791,
780-802,808-850,
836-861,867-920,
906-931,936-978,
964-989,994-1037,
1023-1048,
1053-1096,
1082-1107,
1114-1157,
1143-1168,
1172-1223,
1206-1233
676,721,773,801,
1030,1061,1225
#disulfide_bonds #status predicted\
#binding_site carbohydrate (Asn) (covalent) #status
predicted
SUMMARY #length 1234 #molecular_weight 139081 #checksum 3676
Query Match 84.1%; Score 2834; DB 1; Length 1234;
Best Local Similarity 83.08; Pred. No. 0.00e+00;
Matches 356; Conservative 34; Mismatches 37; Indels 2; Gaps 2;

```

```

Db 199 SNEKRCVEIILCTPPRVNGNGINVKPYKKNERYHYKCKGVYKPERGDAVCTGSGWS 258
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 191 SNEKQCVVEISCLPPRVNGNGIYLKPYKKNERYFOYKCKGFYKERGDAVCTGSGWNP 240
259 QPCEERKRCSPYILNGIYTPHRIHRSDEIRYECNNGFYPVNGSVYKCTPTGMIYVP 318
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 241 QPCEEMVCLLPYIPNGIYTPHRIKRRDDIRYCKNGFYPATRSPVSKCTITGWIYAP 300
319 RCTLKPCFEPQFKGRUYEESLRPNFVPSIGNKYSYKCDNGFSPPS-GYSMDYLCTAQ 377
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 301 RCTLKPCDFPQFKHGRUYEESRRPFPVPKGEYSYCDNGFTTPGQSY-WDYLCTV 359
378 GWEPDVPVRKCVPHYVNGDSAYKGVYVYGOSLKYQCVNGISYLGQDITMTENGS 437
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 360 GWEPDVPCLRCITLHYVEGSSYWRRYEGOSAKYQCHSGYSLPNGODITYTENGS 419
438 PPKCIRIK 446
420 PPKCVIRIK 428
Db 438 PPKCIRIK 446
Qy 420 PPKCVIRIK 428
RESULT 2
ENTRY #type complete
TITLE complement factor H precursor, long splice form - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change
22-Jun-1998
ACCESSIONS S00254; A60238; A54726; A61565; A26505; I72654; S66298
S00254
REFERENCE
#authors Ripoche, J.; Day, A.J.; Harris, T.J.R.; Sim, R.B.
#journal Biochem. J. (1988) 249:593-602
#note The complete amino acid sequence of human complement factor
H.
#cross-references MIMID:88134059
#accession S00254
#molecule_type mRNA
#residues 1-1231 #label RIP
#cross-references EMBL:Y00716; NID:931964; PIDN:CAA68704.1; PID:931965
#note 402-Tyr was also found
parts of this sequence, including the amino and carboxyl
ends of the mature protein, were confirmed by protein
sequencing
REFERENCE
A60238
#authors Estallier, C.; Schwaebler, W.; Dietrich, M.; Weiss, E.H.
#journal Eur. J. Immunol. (1991) 21:799-802
#title Human complement factor H: two factor H proteins are derived
from alternatively spliced transcripts.
#cross-references MIMID:91184292
#accession A60238
#status not compared with conceptual translation
#molecule_type mRNA
#residues 1-56;1177-1231 #label EST
#note only portions of this 4.3 kilobase mRNA were sequenced
A54726
#authors Day, A.J.; Ripoche, J.; Lyons, A.; McIntosh, B.; Harris,
T.J.R.; Sim, R.B.
#journal Biosci. Rep. (1987) 7:201-207
#title Sequence analysis of a cDNA clone encoding the C-terminal end
of human complement factor H.
#cross-references MIMID:88025472
#accession A54726
#status not compared with conceptual translation
#molecule_type mRNA
#residues 'DPRN',579-1231 #label DVY
#cross-references GB:M17517; NID:91804977; PIDN:AAA52076.1; PID:9180498
#note parts of this sequence were determined by protein
sequencing
REFERENCE
A61565
#authors Ripoche, J.; Day, A.J.; Willis, A.C.; Belt, K.T.; Campbell,
R.D.; Sim, R.B.
#journal Biosci. Rep. (1986) 6:65-72
#title Partial characterization of human complement factor H by

```

protein and cDNA sequencing: homology with other complement and non-complement proteins.

##cross-references MUID:86188123

##accession A61565

##status not compared with conceptual translation

##molecule-type mRNA

##residues 1050-1057, 'T', 1059-1102 ##label R12

REFERENCE A26505

##authors Sim, R.B.; Discipio, R.G.

##journal Biochem. J. (1982) 205:285-293

##title Purification and structural studies on the complement-system control protein beta-1-H (factor H).

##cross-references MUID:83048213

##accession A26505

##molecule-type protein

##residues 19-20, 'Q', 22-29, 'V', 31-33, 'Q', 35 ##label SIM

REFERENCE A44551

##authors Barlow, P.N.; Norman, D.G.; Steinkasserer, A.; Horne, T.J.; Pearce, J.; Driscoll, P.C.; Sim, R.B.; Campbell, I.D.

##journal Biochemistry (1992) 31:3626-3634

##title Solution structure of the fifth repeat of factor H: A second example of the complement control protein module.

##cross-references MUID:92232649

##contents annotation: NMR structure determination, residues 264-292

REFERENCE A49224

##authors Norman, D.G.; Barlow, P.N.; Baron, M.; Day, A.J.; Sim, R.B.; Campbell, I.D.

##journal J. Mol. Biol. (1991) 219:717-725

##title Three-dimensional structure of a complement control protein module in solution.

##cross-references MUID:91278097

##contents annotation: NMR structure determination, residues 927-985

REFERENCE 156100

##authors Estaller, C.; Koistinen, V.; Schweabe, W.; Dierich, M.P.; Weiss, E.H.

##journal J. Immunol. (1991) 146:3190-3196

##title Cloning of the 1.4-kb mRNA species of human complement factor H reveals a novel member of the short consensus repeat family related to the carboxy terminal of the classical 150-kDa molecule.

##cross-references MUID:91201892

##accession I72654

##status translated from GB/EMBL/DBJ

##molecule-type mRNA

##residues 1047-1231 ##label RES

##cross-references GB:M65294; NID:g183766; PIDN:AAA35948.1; PID:g183767

REFERENCE S66298

##authors Carton, J.A.; Bates, R.C.; Smith, A.I.; Tetoz, T.; Arellano, A.; Gordon, D.L.; Burns, G.F.

##journal Biochim. Biophys. Acta (1996) 1289:305-311

##title Factor H co-purifies with thrombospondin isolated from platelet secretate.

##cross-references MUID:96205365

##accession S66298

##status preliminary

##molecule-type protein

##residues 411-419, 574-578, 580-582 ##label CAR

COMMENT Factor H has also been found bound to cell membranes in an unknown manner. However, it has at least one cell attachment site motif in repeat 4.

COMMENT Alternative transcripts of 4.3, 1.8, and 1.4 kilobases are expressed in liver. See also PIR:NBH0HS.

GENETICS

##gene GDB:HF1; HF

##cross-references GDB:120041; OMIM:134370

##map_position 1932-1932

GENETICS

##gene GDB:HF2; HF

##cross-references GDB:129095

##map_position 1933-1932

note the correspondence between the two loci and the sequences indicated is unclear; factor H has been reported to have several allelic forms

FUNCTION

#description

a cofactor in the inactivation of C3b by serine proteinase I; also increases the rate of dissociation of the C3bB complex (C3 convertase) and the (C3b)nb complex (C5 convertase) in the alternative complement pathway

CLASSIFICATION

complement alternate pathway

KEYWORDS

superfamily complement factor H; complement factor H repeat

FEATURE

alternative splicing; complement alternate pathway; glycoprotein; plasma

19-1229

19-449

domain signal sequence #status predicted #label SIG

21-80

product complement factor H #status experimental #label MAT

85-141

domain complement factor H repeat homology #label FH01

146-205

domain complement factor H repeat homology #label FH02

210-262

domain complement factor H repeat homology #label FH03

246-248

domain complement factor H repeat homology #label FH04

267-320

region cell attachment (R-G-D) motif

325-385

domain complement factor H repeat homology #label FH05

389-442

domain complement factor H repeat homology #label FH06

448-503

domain complement factor H repeat homology #label FH07

509-564

domain complement factor H repeat homology #label FH08

569-623

domain complement factor H repeat homology #label FH09

630-684

domain complement factor H repeat homology #label FH10

691-744

domain complement factor H repeat homology #label FH11

753-803

domain complement factor H repeat homology #label FH12

811-864

domain complement factor H repeat homology #label FH13

870-926

domain complement factor H repeat homology #label FH14

931-984

domain complement factor H repeat homology #label FH15

989-1043

domain complement factor H repeat homology #label FH16

1048-1102

domain complement factor H repeat homology #label FH17

1109-1163

domain complement factor H repeat homology #label FH18

1167-1228

domain complement factor H repeat homology #label FH19

21-66, 52-80, 85-129,

114-141, 146-192,

178-205, 210-251,

237-262, 267-309,

294-320, 325-374,

357-385, 389-431,

416-442, 448-494,

477-505, 509-553,

536-564, 569-611,

597-623, 630-673,

659-684, 691-733,

719-744, 753-792,

781-803, 811-855,

839-864, 870-915,

901-926, 931-973,

959-984, 989-1032,

1018-1043,

1048-1091,

1077-1102,

1109-1152,

1138-1163,

1167-1218,

1201-1228

217

#disulfide_bonds #status predicted

#binding_site carbohydrate (Asn) (covalent) #status

absent

#binding_site carbohydrate (Asn) (covalent) #status

absent

#binding_site carbohydrate (Asn) (covalent) #status

absent

#binding_site carbohydrate (Asn) (covalent) #status

absent

#binding_site carbohydrate (Asn) (covalent) #status

absent

SUMMARY

#length 1231 #molecular-weight 139124 #checksum 9625

Query Match

Best Local Similarity

Matches

289; Conservative

71.1%;

Score 2396; DB 1;

Length 1231;

Pred. No. 0.00e+00;

Indels 0;

Gaps 0;

Db 19 EDCELPRRNTTEILTGSMDSQTYPEGTQATYKCRPGYSLGNVIMVCRKGEVALNPLR 78
 QY 1 EDCKGPPPRENSSEILSGMSSEOLYSEGTQATYKCRPGYRLGTIVKCKNGEMVPSNSR 60
 Db 79 KCORRPGHPDDTPRGTTTLGNGVFEYGVAVYVNCNGYQLLGEINTRECDTGDWINDI 138
 QY 61 ICRKRPCHPGDTPRGSRRLVSGSEFEFGARVYTCDEGYLLGEIDYRECDADGWINDI 120
 Db 139 PICEVVKCLPTAPENGKIVSSAMEPDRYHFGQAVRVNCSGKRIEEDDEMGHSDGDFW 198
 QY 121 PICEVVKCLPTAPENGKIVSSAMEPDRYHFGQAVRVNCSGKRIEEDDEMGHSDGDFW 180
 Db 199 SKERPKCYEISCKSPDYINGSPISOKIYKKNRPYCKNMNGYSENGDAVCTESGWRP 258
 QY 181 SNEKPPQCVESICLPPRVENGDIYLYKPYKENERFOYKCKGPFYKRGDAVCTESGWRP 240
 Db 259 LPSCSEKSCDPRYPNGDYSPLRIKHRTGDEITQCRNGFPARNGTAKTSGWIPAP 318
 QY 241 QPSCSEMTCLPTYPNGDYSPLRIKHRTGDEITQCRNGFPARNGTAKTSGWIPAP 300
 Db 319 RCTLKPCDYPDIKRGGLYHNMRRPYEPVAVGKYYSYCDHEFTPGSGYWDHICTODG 378
 QY 301 RCTLKPCDYPDIKRGGLYHNMRRPYEPVAVGKYYSYCDHEFTPGSGYWDHICTODG 360
 Db 379 MSPAVPCLRCYFYLNGYVNONHGRKRVGQKSIDVACHGIALPKAQTIVTCMENGWSP 438
 QY 361 WEPEVPLRQCIHYHYEYGESYWRRIEGRQAKVCHSGYSLPNQODTYTCTENGWSP 420
 Db 439 TPRCIRYK 446
 QY 421 PPKCVRIK 428

RESULT 3
 ENTRY NBHUS
 TITLE complement factor H precursor, short splice form - human
 ALTERNATE_NAMES complement factor H-related protein; complement protein H
 ORGANISM Homo sapiens
 DATE 31-Dec-1993 #sequence_revision 23-Feb-1996 #text_change 22-Jun-1999

ACCESSIONS
 REFERENCE S03013; B60238; A27877; A61103; A26505; S10479
 #authors Ripoche, J.; Day, A.J.; Harris, T.J.R.; Sim, R.B.
 #journal Blochem. J. (1988) 249:593-602
 #title The complete amino acid sequence of human complement factor H.

#cross-references MUID:88134059
 #accession S03013
 #molecule_type mRNA
 #residues 1-449 #label RIP
 #cross-references EMBL:X07523; EMBL:Y00716; NID:932492;
 PIDD:CAA30403.1; PIDD:9758073

REFERENCE
 #note part of this sequence, including the amino end of the
 mature protein was confirmed by protein sequencing
 402-Tyr was also found
 A60238

REFERENCE
 #authors Estaller, C.; Schwaible, W.; Dietrich, M.; Weiss, E.H.
 #journal Eur. J. Immunol. (1991) 21:799-802
 #title Human complement factor H: two factor H proteins are derived
 from alternatively spliced transcripts.
 #cross-references MUID:91184292
 #accession B60238
 #status not compared with conceptual translation
 #molecule_type mRNA
 #residues 1-334-449 #label ESR

REFERENCE
 #authors A27877
 #journal Schult, T.F.; Schwaible, W.; Stanley, K.K.; Weiss, E.;
 Dietrich, M.P.
 Eur. J. Immunol. (1986) 16:1351-1355
 #title Human complement factor H: isolation of cDNA clones and
 partial cDNA sequence of the 38-kDa tryptic fragment
 containing the binding site for C3b.

#cross-references MUID:87054207
 #accession A27877
 #molecule_type mRNA

#residues 'IL', 55-401, 'Y', 403-449 #label SCH
 an additional nucleotide present within the codon for
 Glu-310 was thought to be a cloning artifact and was
 ignored in translation

REFERENCE

#authors A61103
 #journal Schwaible, W.; Zwirner, J.; Schulz, T.F.; Linke, R.P.;
 Dietrich, M.P.; Weiss, E.H.
 Eur. J. Immunol. (1987) 17:1485-1489
 #title Human complement factor H: expression of an additional
 truncated gene product of 43 kDa in human liver.
 #cross-references MUID:88055295
 #accession A61103

#status not compared with conceptual translation
 #molecule_type mRNA
 #residues 27-76 #label SC2
 this is a partial sequence of an alternatively spliced
 1.8 kilobase mRNA that is translated to yield a 43 k
 form related to factor H

REFERENCE

#authors A26505
 #journal Sim, R.B.; Discipio, R.G.
 Biochem. J. (1982) 205:285-293
 #title Purification and structural studies on the complement-system
 control protein beta-1-H (factor H).
 #cross-references MUID:83048213
 #accession A26505
 #molecule_type protein

REFERENCE

#authors A44551
 #journal Barlow, P.N.; Norman, D.G.; Steinkasserer, A.; Horne, T.J.;
 Pearce, J.; Driscoll, P.C.; Sim, R.B.; Campbell, I.D.
 Biochemistry (1992) 31:3626-3634
 #title Solution structure of the fifth repeat of factor H: A second
 example of the complement control protein module.
 #cross-references MUID:92322649
 #contents annotation; NMR structure determination, residues 264-292
 S10479

REFERENCE

#authors Kristensen, T.; Wetsel, R.A.; Tack, B.F.
 #journal J. Immunol. (1986) 136:3407-3411
 #title Structural analysis of human complement protein H: homology
 with C4b binding protein, beta(2)-glycoprotein I, and the
 Ba fragment of B.
 #cross-references MUID:86169701
 #accession S10479

#molecule_type mRNA
 #residues 226-401, 'Y', 403-449 #label KRI
 #cross-references GB:M12383; NID:9180472; PIDD:AAA52013.1; PIDD:9180473

COMMENT Factor H has also been found bound to cell membranes in an unknown
 manner. However, it has at least one cell attachment site motif
 in repeat 4.

COMMENT Alternative transcripts of 4.3, 1.8, and 1.4 kilobases are expressed
 in liver. See also PIR:NBHUS.

GENETICS
 #gene GDB:HF1, HF
 #cross-references GDB:120041; OMIM:134370
 #map_position 1q32-1q32

GENETICS
 #gene GDB:HF2, HF
 #cross-references GDB:129095
 #map_position 1q32-1q32

note the correspondence between the two loci and the sequences
 indicated is unclear; factor H has been reported to have
 several allelic forms

FUNCTION

#description a cofactor in the inactivation of C3b by serine proteinase I;
 also increases the rate of dissociation of the C3bb
 complex (C3 convertase) and the (C3b)nb complex (C5
 convertase) in the alternative complement pathway
 #pathway complement alternative pathway
 #classification superfamily complement factor H; complement factor H repeat
 homology

Accession	Gene	Protein	Accession	Gene	Protein
#cross-references MIMD:94318039					
#accession 546199					
##molecule-type mRNA					
##residues 1-1053					
##cross-references EMBL:L21703; NID:g639894; PID:g639895					
##experimental source liver					
#accession 578694					
##molecule-type protein					
##residues 526-532, 'X', 534-537, 809-817, 'X', 819-826					
##label DAH2					
GENETICS					
CLASSIFICATION					
KEYWORDS					
FEATURE					
89-145					
334-389					
450-502					
569-624					
682-738					
743-802					
935-989					
993-1052					
SUMMARY					
Query Match					
Best Local Similarity 31.7%; Pred. No. 1,286-135;					
Matches 133; Conservative 76; Mismatches 180; Indels 28; Gaps 22;					
Db	50	YPGKQVRYGCVGR-S-GFKLYCYEGRN-ETRGAK-CQPSRCGHPGDAQFDFHLAEG 105			
Qy	24	YSEGTQATYKCPGGRITLGIIVKCKNGMVPSPNSPISRIKRRCKGHPGDTFFGSFRLAVG 83			
Db	106	NDFVSGSKVYVYCOGYOWASRINRRCVAEGDGVVPVCSQOC-PLIHVDNNGVYIG 164			
Qy	84	SEFEAGKAVYTCDESGYQLGELIDRECADGTDNDIPICEYVKKCLPYTELENGRIYSGA 143			
Db	165	--PEEAT-FGNVYRFSCKSRSEIILDGSPELCYCDERDGSVPYKCKAITCAIPIENGAV 221			
Qy	144	AEPDQEVYFGVYVRECNGSFKI-EGQKEMHCSENGLSMSEKPOQCEISCLPREEVNGDG 202			
Db	222	PGAIKEXKENDVILHEDCAFHKHIDRPSICIKGIAEWSPPPLCSIKICRLTIMDGT 281			
Qy	203	YLKPYVKNERNQYKCKGGEVYKKEGDA-VCTG--SGNNPDSCEMTCLTPYIPNGIT 259			
Db	282	EPAYRNLFSPEGLTKVICTARTSWISTPOETSVTYTCCODNGEWSIRPTOCYERCSNRPEH 341			
Qy	260	TP-HRIKRIHIDEIRVEG-KNGFYPAIR-SPV-SKCLITG-WIPAPRCSLKPDQPFQKH 314			
Db	342	VDSMDVRSERY--TLDDNTRYCKRGYKRTGGVTA--TCGRNGMPPPLEYVTKCSK 396			
Qy	315	GRLYYEESRRPFPVPDYGKEYSYCCNGFTTPSOASYWDYLRCTVNGEPEVPC-LRQCIF 373			
Db	397	ENIDQAVIYGTGKOIYNLQKAIYACGEG-N--RGRITTLGCGNSGDKRCT-VK 448			
Qy	374	HYVEGESSYWMQR-YIEGOSAKVQCHSGVSLNQGDTTYCTENGWSPPPKCYARIK 428			
RESULT	9				
ENTRY	A45222	#type complete			
TITLE		complement factor H-related protein DOWN16 precursor - human			
ORGANISM		#formal name Homo sapiens #common name man			
DATE	30-Sep-1993	#sequence-revision 30-Sep-1993			
ACCESSIONS	24-Sep-1998				
REFERENCE	A45222; S29609				
#authors	A45222				
	Stekler, C.; Kuhn, S.; Gunther, K.; Lingelbach, K.; Zipfel, P.F.				
#journal	J. Biol. Chem. (1993) 268:2904-2908				
#title	A novel short consensus repeat-containing molecule is related to human complement factor H.				
#cross-references MIMD:93153112					

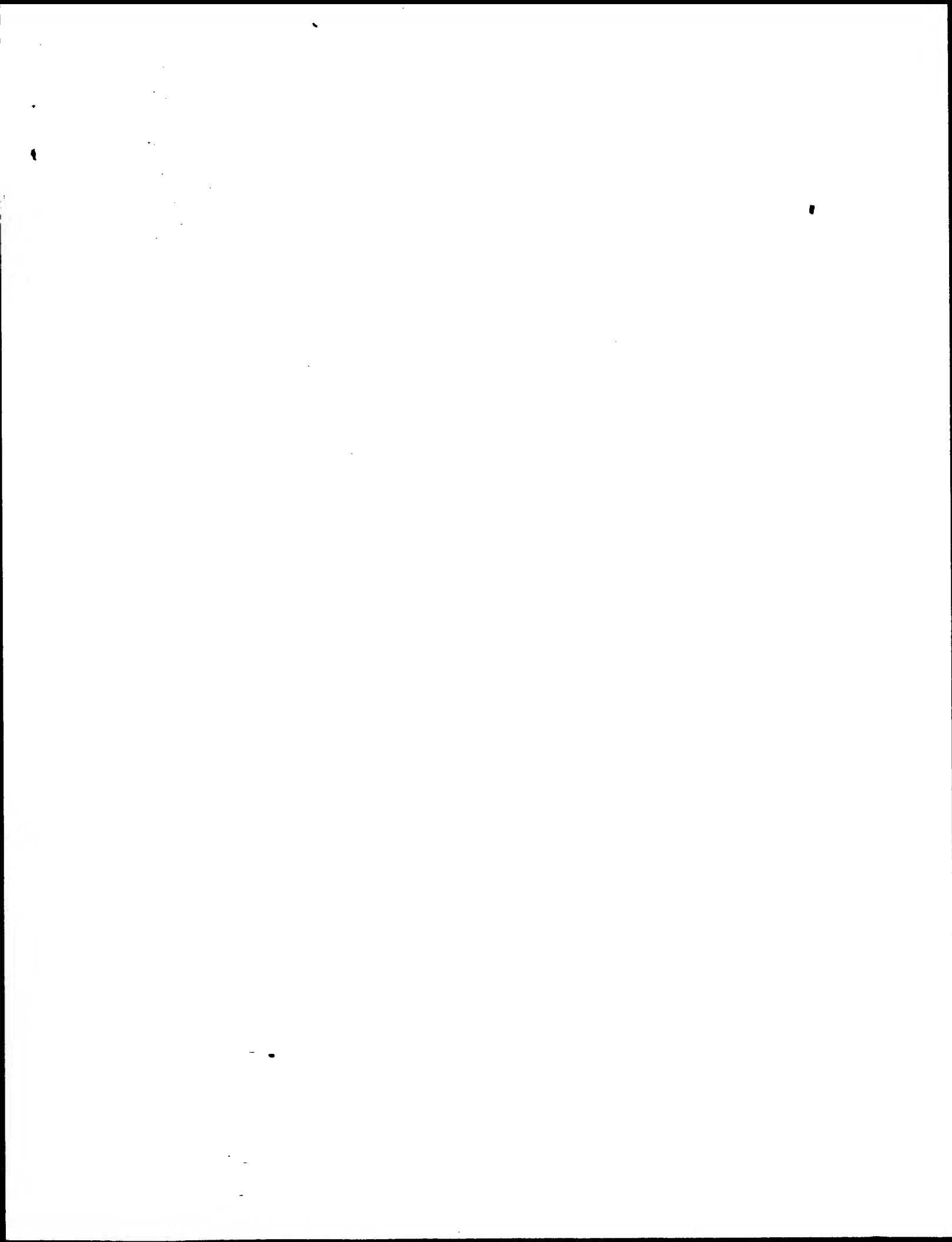
#accession	A45222
#molecule_type	mRNA
#residues	1-331 ##label SRK
#cross-references EMBL:X68679;	NID:g30869; PTD:g30870
#experimental_source liver	
#note	sequence extracted from NCBI backbone (NCBIP:124269)
GENETICS	
#gene	GDB:HPL4
#cross-references GDB:137201	
CLASSIFICATION	#superfamily complement factor H repeat homology
FEATURES	glycoprotein
FEATURE	
1-18	
19-331	
23-83	
87-141	
147-204	
211-265	
269-330	
108,186,206,310	
SUMMARY	
#length 331	#molecular-weight 37469 #checksum 6811
Query Match	19.2%; Score 646; DB 2; Length 331;
Best Local Similarity	56.3%; Pred. No.2,32e+126;
Matches	76; Conservative 26; Mismatches 32; Indels 1; Gaps 1;
Db	9 LTLWSCANQVPCFPDPIKHGFLFHEMMRRPVPYAVAKIYSYCDEHFTPSGYWD 68 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : QY 293 ITGMIRAPROSLRKCDDPEQRKHRLKYESSRRRYEPVPIGKIYSYCDGDGFPTPSQSYWD 352
Db	69 YIHQTONGSPAYPCLRKCYTPLYENGYNQNTGRKVQGNSTEVACHRGYLPRKRTTY 128 : : : : : : : : : : : : : : : : : : : : : : : : : QY 353 YLRCTVGWSPEPCELCROCIFHYEYGESSYWQORYIEQSARVAOCHSGSLPNQ-IDTY 411
Db	129 TCTENGMSPPRCIR 143 : : : : : : : : : : : : : : : : : : : : : : : QY 412 YCTENGMSPPCKVR 426
ENTRY	RESULT 10
TITLE	S53711 #type complete
ABSTRACT	C4BP alpha chain precursor - rabbit
ORGANISM	#formal_name Oryctolagus cuniculus #common_name domestic #rabbit
DATE	10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
ACCESSIONS	S53711
REFERENCE	S53711
#authors	de Frutos, P.G.; Dahlbaeck, B.
#journal	Biochim. Biophys. Acta (1995) 1261:285-289
#title	CDNA Structure of rabbit C4b-binding protein alpha-chain. Preserved sequence motif in complement regulatory protein modules which bind C4b.
#cross-references MUID:95226458	
#accession	S53711
#status	Preliminary: nucleic acid sequence not shown
#molecule_type mRNA	
#residues	1-597 ##label DEF
#cross-references EMBL:I235490	
CLASSIFICATION	#superfamily C4b-binding protein alpha chain; complement factor H repeat homology
FEATURE	
50-107	
112-169	
174-234	
239-294	
299-360	
364-422	
426-480	
484-538	
SUMMARY	#length 597 #molecular-weight 66130 #checksum 6473


```

#authors      Aso, T.; Okamura, S.; Matsuguchi, T.; Sakamoto, N.; Sata, T.,
#journal      Nihon. Y. Biochem. Biophys. Res. Commun. (1991) 174:222-227
#title        Genomic organization of the alpha chain of the human
#cross-references MIMID:91113199
#accession    I52244
##status      translation not shown; translated from GB/EMBL/DBJ
##molecule-type DNA
##residues    1-597 ##label ASO
##cross-references GB:M62486; NID:9190498; PIDN:AAA36506.1; PID:9190500
COMMENT       C4BP controls the classical pathway of complement activation. It
              binds as a cofactor to C3b/C4b inactivator (C3bINA), which then
              hydrolyzes the complement fragment C4b. It also accelerates the
              degradation of the C4bC2a complex (C3 convertase) by dissociating
              the complement fragment C2a.
COMMENT       C4BP occurs in plasma in two forms, both of which bind complement
              fragment C4b. The prevalent higher molecular weight form contains
              7 alpha chains and one beta chain, which are linked by disulfide
              bonds. The beta chain binds the vitamin K-dependent plasma
              protein S. A minor form lacks the beta chain. Bound protein S is
              inactive as a cofactor for protein C inactivation of coagulation
              factors V and VIII.
COMMENT       The molecule has a central body supporting seven tentacles (alpha
              chains), each with the binding site for C4b at the peripheral
              end.
GENETICS
#gene         GDB:C4BPA
##cross-references GDB:120568; OMIM:120830
#map_position 1932-1932
#introns      48/1; 110/1; 143/2; 172/1; 236/1; 297/1; 362/1; 425/1; 482/1;
              540/3
COMPLEX
#classification octamer of seven alpha chains and one beta chain
#superfamily    C4b-binding protein alpha chain; complement
              factor H repeat homology
KEYWORDS       acute phase; chylomicron; complement pathway; duplication;
              glycoprotein; plasma
FEATURE
#feature        #domain signal sequence #status predicted #label SIG\
1-48            #product C4b-binding protein alpha chain #status
              predicted #label MAT\
49-597          #domain complement factor H repeat homology #label FH1\
              #domain complement factor H repeat homology #label FH2\
              #domain complement factor H repeat homology #label FH3\
              #domain complement factor H repeat homology #label FH4\
              #domain complement factor H repeat homology #label FH5\
              #domain complement factor H repeat homology #label FH6\
              #region complement C4b binding #status predicted\
              #domain complement factor H repeat homology #label FH7\
              #domain complement factor H repeat homology #label FH8\
              #binding_site carboxydrate (Asn) (covalent) #status
              experimental
SUMMARY        #length 597 #molecular-weight 67033 #checksum 6374
Query Match    15.9%; Score 537; DB 1; Length 597;
Best Local Similarity 30.7%; Pred. No. 2,52e-99;
Matches 138; Conservative 85; Mismatches 182; Indels 45; Gaps 40;
Db 49 NC-GPPTLSAAMDITLRETRKCTTLKYCLGCVYVSHSTQTLTNSGSEW-VYT 106
QY 2 DCKGPPREN-SEILSGWSBOLTBETQATYKCRPQY-TTLTIVKYVK-NGEWPSNP 58
Db 107 FCII-KR-CRHPGELRNG--OVEIKTDLTSGSIOIEFCSGEPFLISSTSRCEVDORGV 162
QY 59 SRICKRRCGHPGDPPTGSPFLANVGSFEFGAKVYVTCDDGYLLDEIDYR-BC-DAD-G 115
Db 163 WSHPLQCEIVKXKPPDIRNGR-HSG-E-ENFYAYGFSVYSCDPRSLGHASISCT 218
QY 116 WTNDPIQCEVVKCLPVELENGRIVSAAPDEQYVGVVRECNCSGFIKQKEMHCS 175
Db 219 VENEIIGVWRSPPCKEITCRKDVSHGEMVSGFGIYVYKTIYFKQKGVTLRGSSV 278
QY 176 -EN--GLMSNEKQOCVEISCLPRLVYNGGII-LKRVYKENERFYQKKGQVYKERD 230

```

Db	219	ICHOADSKMNSPPRACENPCNINCPDIHASMWEIYPRPTEDVYVGVLYRNCYCHGKXP	338
Qy	231	AVC-TGSGWNP-QSCSEMTCLT-PLYNPGLY-T-PIRINHRI--DDEI-RYCKNGEYF	282
Db	339	TTDEPTVYICKNLRMPYOGCEALCCPEKLNNGEIT-QH-EKS-RPANHC-VY-FYGD	393
Qy	283	ATRSPVY-KCIT-T-GMIPAPRCLAKPDFQFKHGRLYEESNRPFPPVIGKEYSYCD	340
Db	394	EISFCHETSFEFSAI-CQGDGTMSPPRPGCDICNP-PIAHGHTKQSSSYFFFEETI	451
Qy	341	N-GFTTSQSYWDYLRCTVNG-WEPEVP-CLROCFHYEYVEGSSYWRQ-RYEGOSAKV	396
Db	452	YECDKGYLLV-GOAKKSCSYSHMSVAPROC	480
Qy	397	-QCHSGISLPNGQDTTYCTENGWSP-PRC	424
RESULT 12			
ENTRY	S57953 #type complete		
TITLE	C4B protein alpha chain precursor - rat		
ORGANISM	#formal_name Rattus norvegicus #common_name Norway rat		
DATE	13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Aug-1999		
ACCESSIONS	S57953		
REFERENCE	S57953		
#authors	Hillarp, A.; Thern, A.; Dahlback, B.		
#submission	Submitted to the EMBL Data Library, July 1995		
#description	Molecular cloning of rat C4b-binding protein alpha- and beta-chains: structural and functional relationships among human, bovine, rabbit, mouse and rat proteins.		
#accession	S57953		
#status	preliminary		
#molecule_type	mRNA		
#residues	1-558 #label HIL		
#cross-references	EMBL:Z50051; NID:g899379; PIDN:CAA9391.1; PID:g899380		
CLASSIFICATION	#superfamily C4b-binding protein alpha chain; complement factor H repeat homology		
FEATURE			
15-72	#domain complement factor H repeat homology #label FH1		
77-134	#domain complement factor H repeat homology #label FH2		
139-199	#domain complement factor H repeat homology #label FH3		
204-258	#domain complement factor H repeat homology #label FH4		
263-324	#domain complement factor H repeat homology #label FH5		
328-386	#domain complement factor H repeat homology #label FH6		
390-443	#domain complement factor H repeat homology #label FH7		
447-501	#domain complement factor H repeat homology #label FH8		
SUMMARY	#length 558 #molecular_weight 62465 #checksum 9928		
Query Match	14.7%: Score 494; DB 2; Length 558;		
Best Local Similarity	29.5%; Pred. No. 8.78e-89;		
Matches 132; Conservative	89; Mismatches 182; Indels 45; Gaps 38;		
Db	15	C-GPPP-DLPALPASENNQTDIFESHITLRLKNCPRGYSRRASSQSOLYCKPLGKW-QINIA	71
Qy	3	CKGPPRENSIELSGSWEQI-YSEGQATATYKCRGY-RTIGTIVYCKN-GEWVPENPS	59
Db	72	--CVAKSCRNPDQIANG--KYEVKTDLFQSGIPESSCEGIIIGSSTSCIEIOGKGVSW	127
Qy	60	RICRRRPPCGHGDDPFGFSFRILAVSEEFEGAKVYITDCEGYOLGGE-IDYRECDAGD--W	116
Db	128	SDPLRECIYACGMW-PIISNGK-HNGC-E-EEFFTRSSVYTKCDDDFLLGASITCT	183
Qy	117	TNDIPICEVAKC-LPVTLENGRIYGAALPDQETGYGVYRFCSNGSKFIEGKEMHKS	175
Db	184	VYVKTGVWSPSPPTCEHIIICPMWRVYLGHTINSFKAHTYKDSVRFVYCKGFLRSGV	243
Qy	176	E-N--GLMSNKKPCVCVETICLIPRENGD-GIYLKPYKENEERFQYKCKGQGFYKRGD	230
Db	244	IHC6ADSGMSPPVCEVLSNCDIDIPDIAALITSPRPRKEVYVGVVLYRICRGTPEPA	303
Qy	231	AVCTSGG-WNDPSCSEETCL-TPYIPNG--IYHP-RIKH-RIDELIREKNGEYPA	283



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

```

Title: >US-09-316-163-14
Description: (1-428) from US09316163.pep
Perfect Score: 3371
Sequence: 1 EDCCGPPPPRENSSELLSSWS.....DTYCTENGMSPPKCVRIK 428

```

Scoring table: PAM 150

Searched: 83857 seqs, 30454973 residues

Listing first 45 summaries

Database: swiss-prot38

Statistics: Mean ~~46.246~~; Variance 66.763; scale 0.693

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	2834	84.1	1234	1	CEAH_MOUSE	COMPLEMENT FACTOR H PR	0.00e+00
2	2396	71.1	1231	1	CEAH_HUMAN	COMPLEMENT FACTOR H PR	0.00e+00
3	646	19.2	331	1	CFHD_HUMAN	COMPLEMENT FACTOR H-LI	8.26e-14
4	537	15.9	357	1	CFBD_HUMAN	CAB-BINDING PROTEIN AL	1.67e-11
5	494	14.7	558	1	CABP_RAT	CAB-BINDING PROTEIN AL	1.09e-9
6	474	14.1	610	1	CABP_BOVIN	CAB-BINDING PROTEIN AL	3.15e-9
7	446	13.2	2039	1	CRI_HUMAN	COMPLEMENT RECEPTOR TY	1.27e-8
8	445	12.9	330	1	CEH1_HUMAN	COMPLEMENT FACTOR H-LI	2.23e-8
9	434	12.9	468	1	CABP_MOUSE	CAB-BINDING PROTEIN PR	2.23e-8
10	431	12.9	270	1	CFHE_HUMAN	COMPLEMENT FACTOR H-LI	1.44e-8
11	396	11.7	263	1	VCP_VACCV	COMPLEMENT CONTROL PRO	3.54e-7
12	390	11.6	830	1	LEM3_HUMAN	P-SELECTIN PRECURSOR (1.41e-7
13	375	11.1	345	1	APOH_MOUSE	BETA-2-GLYCOPROTEIN I	1.37e-6
14	374	11.1	768	1	LEM3_RAT	P-SELECTIN PRECURSOR (1.51e-6
15	367	10.9	345	1	APOH_HUMAN	BETA-2-GLYCOPROTEIN I	1.79e-6
16	368	10.9	768	1	LEM3_MOUSE	P-SELECTIN PRECURSOR (9.73e-6
17	364	10.8	345	1	APOH_CANFA	BETA-2-GLYCOPROTEIN I	1.11e-6
18	357	10.6	610	1	LEM2_HUMAN	E-SELECTIN PRECURSOR (7.76e-6
19	359	10.6	612	1	LEM2_MOUSE	E-SELECTIN PRECURSOR (2.31e-6
20	348	10.3	668	1	FL3B_MOUSE	COAGULATION FACTOR XII	1.80e-6
21	345	10.2	377	1	MCP_HUMAN	MEMBRANE COFACTOR PROT	1.10e-5
22	344	10.2	1025	1	CR2_MOUSE	COMPLEMENT RECEPTOR TY	2.01e-5
23	341	10.1	345	1	APOH_BOVIN	BETA-2-GLYCOPROTEIN I	1.23e-5

24	335	9.9	759	1	LEM3_SHEEP	P-SELECTIN PRECURSOR (4.52e-55
25	329	9.8	551	1	LEM3_RABBIT	E-SELECTIN PRECURSOR (1.65e-55
26	338	9.7	507	1	DAF_CANARO	COMPLEMENT DECAY-ACCEL	3.00e-55
27	327	9.7	661	1	F13B_HUMAN	COAGULATION-RELATED XIT	5.47e-55
28	328	9.7	978	1	HIG_DROME	LOCOMOTION-RELATED PRO	3.00e-55
29	323	9.6	484	1	LEM3_PIG	E-SELECTIN PRECURSOR (5.97e-54
30	313	9.3	381	1	DAF_HUMAN	COMPLEMENT DECAY-ACCEL	2.30e-51
31	309	9.2	297	1	APOL_RAT	BETA-2-GLYCOPROTEIN I	2.47e-50
32	306	9.1	1033	1	CR2_HUMAN	COMPLEMENT RECEPTOR TY	1.46e-49
33	332	9.0	611	1	LEM2_CANFA	E-SELECTIN PRECURSOR (4.77e-49
34	301	8.9	485	1	LEM2_BOVIN	E-SELECTIN PRECURSOR (2.81e-48
35	294	8.7	340	1	DAF_PONDY	COMPLEMENT DECAY-ACCEL	1.73e-46
36	292	8.7	549	1	LEM3_RAT	E-SELECTIN PRECURSOR (5.61e-46
37	291	8.6	390	1	DAF1_MOUSE	COMPLEMENT DECAY-ACCEL	1.01e-45
38	276	8.2	407	1	DAF2_MOUSE	COMPLEMENT DECAY-ACCEL	6.42e-42
39	274	8.1	360	1	CCP6_HYSEA	COMPLEMENT CONTROL PRO	2.05e-41
40	268	8.0	646	1	LEM3_BOVIN	P-SELECTIN PRECURSOR (6.51e-40
41	255	7.6	1019	1	LEF3_TACIR	LIMULUS CLOTTING FACTO	1.16e-36
42	232	6.9	372	1	LEM1_MOUSE	L-SELECTIN PRECURSOR (5.24e-31
43	226	6.7	372	1	LEM1_PAPHA	L-SELECTIN PRECURSOR (1.49e-29
44	226	6.7	372	1	LEM1_MACMU	L-SELECTIN PRECURSOR (1.49e-29
45	225	6.7	372	1	LEM1_RAT	L-SELECTIN PRECURSOR (2.60e-29

ALIGNMENTS

```

CC RESULT 1
CC ID CFAH_MOUSE STANDARD; PRT: 1234 AA.
CC AC P06909;
CC DT 01-JAN-1988 (Rel. 06, Created)
CC DT 01-JAN-1988 (Rel. 06, Last sequence update)
CC DT 01-NOV-1997 (Rel. 35, Last annotation update)
CC DE COMPLEMENT FACTOR H PRECURSOR (PROTEIN BETA-1-H).
CC GN CFH.
CC OS Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
CC OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC RN
CC RP [1]
CC RX SEQUENCE FROM N.A.
CC RX MEDLINE; 86233353.
CC RA Kristensen T., Tack B.F.;
CC RT "Murine protein H is comprised of 20 repeating units, 61 amino acids
CC RT in length.";
CC RL Proc. Natl. Acad. Sci. U.S.A. 83:3963-3967(1986).
CC RN [2]
CC RP SEQUENCE OF 1-19 FROM N.A.
CC RC STRAIN-BALB/C;
CC RX MEDLINE; 90148935.
CC RA Munoz-Canoes P., Tack B.F., Yik D.P.;
CC RT "Analysis of complement factor H mRNA expression: dexamethasone and
CC RT IFN-gamma increase the level of H in L cells.";
CC RL Biochemistry 28:9891-9897(1989).
CC RN [3]
CC RP SEQUENCE OF 1-18 FROM N.A.
CC RX MEDLINE; 90111033.
CC RA Natunue-sakai S., Nonaka M., Harada Y.N., Shreffler D.C.,
CC RA Moriwaiki K.;
CC RT "Demonstration of an unusual allelic variation of mouse factor H by
CC RT the complete cDNA sequence of the H.2 allotype.";
CC RL J. Immunol. 144:358-362(1990).
CC CC -I- FUNCTION: FACTOR H FUNCTIONS AS A COFACTOR IN THE INACTIVATION OF
CC CC C3B BY FACTOR I AND ALSO INCREASES THE RATE OF DISSOCIATION OF THE
CC CC C3BB COMPLEX (C3 CONVERTASE) AND THE (C3)NB COMPLEX (C5
CC CC CONVERTASE) IN THE ALTERNATIVE COMPLEMENT PATHWAY.
CC CC -I- POLYMORPHISM: TWO CODOMINANT ALLELES OF FACTOR H ARE PRESENT IN
CC CC MICE.
CC CC -I- SIMILARITY: CONTAINS 20 SUSHI (SCR) REPEATS.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial

```



```
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL; M12660; AAA37759.1; -
DR EMBL; J02891; AAA37795.1; -
DR EMBL; M31979; AAA37762.1; -
DR PIR; A26154; NEMSH.
DR HSSP; P08603; 1HFI.
DR MGD; MGI:88385; CFH.
DR PRAM; PFO0084; sushi; 20.
KM Complement alternate pathway: Plasma; Glycoprotein; Repeat; Sushi;
KM Signal.
FT SIGNAL. 1 18 BY SIMILARITY.
FT CHAIN 19 1234 COMPLEMENT FACTOR H.
FT DOMAIN 20 1224 20 x SUSHI (SCR) REPEATS.
FT REPEAT 20 81 SUSHI 1.
FT REPEAT 84 142 SUSHI 2.
FT REPEAT 145 206 SUSHI 3.
FT REPEAT 209 263 SUSHI 4.
FT REPEAT 266 321 SUSHI 5.
FT REPEAT 324 386 SUSHI 6.
FT REPEAT 388 443 SUSHI 7.
FT REPEAT 447 506 SUSHI 8.
FT REPEAT 508 565 SUSHI 9.
FT REPEAT 568 623 SUSHI 10.
FT REPEAT 628 684 SUSHI 11.
FT REPEAT 744 803 SUSHI 12.
FT REPEAT 807 862 SUSHI 13.
FT REPEAT 866 932 SUSHI 14.
FT REPEAT 935 990 SUSHI 15.
FT REPEAT 993 1049 SUSHI 16.
FT REPEAT 1052 1108 SUSHI 17.
FT REPEAT 1113 1169 SUSHI 18.
FT REPEAT 1171 1234 SUSHI 19.
FT DISULFID 21 66 BY SIMILARITY.
FT DISULFID 52 80 BY SIMILARITY.
FT DISULFID 85 129 BY SIMILARITY.
FT DISULFID 114 141 BY SIMILARITY.
FT DISULFID 146 192 BY SIMILARITY.
FT DISULFID 178 205 BY SIMILARITY.
FT DISULFID 210 251 BY SIMILARITY.
FT DISULFID 237 262 BY SIMILARITY.
FT DISULFID 267 309 BY SIMILARITY.
FT DISULFID 284 320 BY SIMILARITY.
FT DISULFID 325 374 BY SIMILARITY.
FT DISULFID 357 385 BY SIMILARITY.
FT DISULFID 389 431 BY SIMILARITY.
FT DISULFID 416 442 BY SIMILARITY.
FT DISULFID 448 494 BY SIMILARITY.
FT DISULFID 477 505 BY SIMILARITY.
FT DISULFID 509 553 BY SIMILARITY.
FT DISULFID 536 564 BY SIMILARITY.
FT DISULFID 569 610 BY SIMILARITY.
FT DISULFID 597 622 BY SIMILARITY.
FT DISULFID 629 672 BY SIMILARITY.
FT DISULFID 658 683 BY SIMILARITY.
FT DISULFID 690 732 BY SIMILARITY.
FT DISULFID 718 743 BY SIMILARITY.
FT DISULFID 732 791 BY SIMILARITY.
FT DISULFID 780 802 BY SIMILARITY.
FT DISULFID 808 850 BY SIMILARITY.
FT DISULFID 836 861 BY SIMILARITY.
FT DISULFID 867 920 BY SIMILARITY.
FT DISULFID 906 931 BY SIMILARITY.
FT DISULFID 936 978 BY SIMILARITY.
FT DISULFID 964 989 BY SIMILARITY.
FT DISULFID 994 1037 BY SIMILARITY.
FT DISULFID 1023 1048 BY SIMILARITY.
FT DISULFID 1053 1096 BY SIMILARITY.
FT DISULFID 1082 1107 BY SIMILARITY.
FT DISULFID 1114 1157 BY SIMILARITY.
FT DISULFID 1143 1168 BY SIMILARITY.
```

FT	DISULFID	1172	1223	BY SIMILARITY.	
FT <td>DISULFID</td> <td>1206</td> <td>1233</td> <td>BY SIMILARITY.</td>	DISULFID	1206	1233	BY SIMILARITY.	
FT	CARBOHYD	676	676	POTENTIAL.	
FT	CARBOHYD	721	721	POTENTIAL.	
FT	CARBOHYD	773	773	POTENTIAL.	
FT	CARBOHYD	801	801	POTENTIAL.	
FT	CARBOHYD	1030	1030	POTENTIAL.	
FT	CARBOHYD	1061	1061	POTENTIAL.	
FT	CARBOHYD	1225	1225	POTENTIAL.	
SQ	SEQUENCE	1234 AA;	139082 MM;	C5AC02P341B95/F7 CRC64;	
Query Match					
Best Local Similarity		84.1%;	Score 2834;	DB 1; Length 1234;	
Matches 356;		Conservative	34;	Mismatches 37; Indels 2; Gaps	
Db	19	EDCKGPPRENEILLGSWS	EDLYEDGQATYKCKPGRITLGT	IVYKCKNGKWA	NSPSR 78
QY	1	EDCKGPPRENEILLGSWS	EQLYSEGTQATYKCRPGYRTLT	ITVYKCKNGEWP	NSPSR 60
Db	79	ICRKKPCGHPGDTPGSF	RLAVAGSOFEEGAKVYTCDDG	YOLGELIDYREGAG	WIMDI 13
QY	61	ICRKRPCGHPGDTPGSF	RLAVAGSFEFGAKVYTCDDG	YOLGELIDYREGAD	WIMDI 12
Db	139	PLCEVYVKKCLPTLE	NGRIVSGAETDQYVYFGVY	VFECNSGFKIEGKE	HICSEGLW 19
QY	121	PLCEVYVKKCLPTLE	NGRIVSGAETDQYVYFGVY	VFECNSGFKIEGK	HICSEGLW 18
Db	199	SNEKRCVEILCTP	PERVENGDGINKPYKKNERYH	YCKGKGVYKPKERGA	VACTGSGMS 25
QY	181	SNEKRCVEILCTP	PERVENGDGINKPYKKNERYH	YCKGKGVYKPKERGA	VACTGSGMS 24
Db	259	QPFCEKRCSPYIL	INGIYTPHRIIHRSDDEIR	ECNNGFYVVGSTVSK	CTPGWIIPV 31
QY	241	QPSCEMCLTPY	INGIYTPHRIIHRSDDEIR	ECNNGFYVVGSTVSK	CTPGWIIPV 30
Db	319	RCILKPCGEPCKYGR	LYEESLRPNFVSYIGNY	SYKCKDNGSEPPS	-GISWDLRCTAQ 37
QY	301	RCSLPCGEPCKYGR	LYEESLRPNFVSYIGNY	SYKCKDNGSEPPS	-GISWDLRCTAQ 35
Db	378	GMEPEVPCVRKCV	HYVENGDSAAWEKYYVQ	GSGLKQCYNGSYL	ONGODMTCTENGMS 43
QY	360	GMEPEVPCVRKCV	HYVENGDSAAWEKYYVQ	GSGLKQCYNGSYL	ONGODMTCTENGMS 41
Db	438	PPPKCIRIK	446		
QY	420	PPPKCIRIK	428		
RESULT 2 STANDARD; PRT: 1231 AA.					
ID	CEAH_HUMAN				
AC	P08603;				
DT	01-AUG-1988 (Rel. 08, Created)				
DT	01-JAN-1990 (Rel. 13, Last sequence update)				
DT	15-FEB-2000 (Rel. 39, Last annotation update)				
DE	COMPLEMENT FACTOR H PRECURSOR.				
GN	HF1 OR HF OR CEF.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
OC	Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RP	TISSUE=LIVER.				
RX	MEDLINE: 88134059.				
RA	Ripoche J., Day A.J., Harris T.J.R., Sim R.B.;				
RU	"The complete amino acid sequence of human complement factor H";				
RT	Biochem. J. 249:593-602(1988).				
RT	[2]				
RP	SEQUENCE OF 53-445 FROM N.A.				
RX	MEDLINE: 87054207.				
RA	Schulz T.F., Schwabe W., Stanley K.K., Weiss E., Dierich M.P.;				
RT	"Human complement factor H: Isolation of cDNA clones and partial cDNA				
RT	sequence of the 36-kDa tryptic fragment containing the binding site				
RT	for C3b.";				

FT	VARIANT	402	402	H -> Y.
FT	CONFLICT	21	21	/FTIG=VAR_001979.
FT	CONFLICT	30	30	C -> Q (IN REF. 3).
FT	CONFLICT	34	34	T -> V (IN REF. 3).
FT	CONFLICT	53	54	T -> Q (IN REF. 3).
FT	STRAND	870	871	RP -> IL (IN REF. 2).
FT	STRAND	876	876	
FT	TURN	877	877	
FT	STRAND	879	881	
FT	STRAND	890	891	
FT	TURN	893	894	
FT	STRAND	896	900	
FT	TURN	903	904	
FT	STRAND	907	907	
FT	STRAND	912	916	
FT	TURN	917	918	
FT	STRAND	919	920	
FT	STRAND	926	926	
SO	SEQUENCE	1231 AA:	123125 MW:	C65DEC8CF8800B3FD CRC64;

DB 219 VENETIGVWNPSPPTCEKTCRKPDVSHGEMVSGFPIYNYKTIIVKCOKEFLRGSSV 278
 OY 176 -EN---GLMSNEKQCVETISCLPPRVENGSDI-LKPYKNEFQYKCKOGFYKKERGD 230
 DB 279 IHCDAOSKNMSPSPACPNCSINLPDIPLHASMETYPPTEDVYVGVTLRYRCHPGYKP 338
 OY 231 AVCTGSGWNP-QPSCHEMTCLT-PIYIPNGIY-T-PRIRKHRI--DDEL-RYECKNGFYF 282
 DB 339 TIDEPTVICOKNLRTMPTPYGCEALCCPEKKNNGEIT-OH-RKS-RPANHC-VY-FYCD 393
 OY 283 ATSPYS-KCTIT-GWIPARCSLPCDFQPFHGRGLYEESSRPFPVPDYGKESYCD 340
 DB 394 EISFSCHEISFSAI-CQGDGTWSPRTSPGCDICNFP-PIAHGHYKSSSYSPFKEEII 451
 OY 341 N-GFTTSPQSYMDIYCTVNG-WEPEVP-CLROCIPIHYVEGESSYWMQR-YIEGOSAKV 396
 DB 452 YECDKGYILV-GOAKLSCSYSHWSAPAPQC 480
 OY 397 -OCHSGYSLPNGODTYCTENGWSP-PRC 424

RESULT 5
 ID C4BP_RAT STANDARD; PRT; 558 AA.
 AC 063514;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE C4B-BINDING PROTEIN ALPHA CHAIN PRECURSOR.
 GN C4BPA.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN:SPRAGUE-DAWLEY; TISSUE:LIVER;
 RX MEDLINE: 97166082.
 RA Hillarp A., Wiklund H., Thern A., Dahlback B.,
 RT "Molecular cloning of rat C4b binding protein alpha- and beta-chains:
 RT structural and functional relationships among human, bovine, rabbit,
 RT mouse, and rat proteins.";
 RL J. Immunol. 158:1315-1323(1997).
 CC -1- FUNCTION: C4BP CONTROLS THE CLASSICAL PATHWAY OF COMPLEMENT
 CC ACTIVATION. IT BINDS AS A COFACTOR TO C3B/C4B INACTIVATOR
 CC (C3BINA), WHICH THEN HYDROLYSES THE COMPLEMENT FRAGMENT C4B. IT
 CC ALSO ACCELERATES THE DEGRADATION OF THE C4B C2A COMPLEX (C3
 CC CONVERTASE) BY DISSOCIATING THE COMPLEMENT FRAGMENT C2A. ALPHA
 CC CHAIN BINDS C4B. IT INTERACTS ALSO WITH ANTICOAGULANT PROTEIN S
 CC AND WITH SERUM AMYLOID P COMPONENT.
 CC -1- SUBUNIT: DISULFIDE-LINKED COMPLEX OF C4BP ALPHA AND BETA CHAINS.
 CC -1- SIMILARITY: TO C4BP BETA CHAIN AND TO PIG APOLIPOPROTEIN R.
 CC -1- SIMILARITY: CONTAINS 8 SUSHI (SCR) REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by, and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z50051; CAA90391.1; -.
 DR HSSP: P10998; IYVC.
 DR PFM: PFO0084; SUSHI; 8.
 KW Complement pathway; Plasma; Glycoprotein; Repeat; Sushi; Signal.
 FT CHAIN 1 13
 FT SIGNAL 13
 FT DOMAIN 14 558
 FT REPEAT 14 502
 FT REPEAT 14 73
 FT REPEAT 14 76
 FT REPEAT 138 200
 FT REPEAT 203 259
 FT REPEAT 262 325
 FT SUSHI 5.

FT REPEAT 328 387
 FT REPEAT 389 444
 FT REPEAT 446 502
 FT DISULFID 15 60
 FT DISULFID 45 72
 FT DISULFID 77 118
 FT DISULFID 104 134
 FT DISULFID 139 182
 FT DISULFID 168 199
 FT DISULFID 204 246
 FT DISULFID 232 258
 FT DISULFID 263 312
 FT DISULFID 296 324
 FT DISULFID 328 351
 FT DISULFID 329 373
 FT DISULFID 363 386
 FT DISULFID 390 431
 FT DISULFID 417 443
 FT DISULFID 447 488
 FT DISULFID 474 501
 FT DISULFID 509 509
 FT DISULFID 521 521
 FT CARBOHYD 31 31
 FT CARBOHYD 177 177
 FT CARBOHYD 186 186
 FT CARBOHYD 469 469
 FT CARBOHYD 491 491
 SQ SEQUENCE 558 AA; 62266 MW; 592F0C67ED1E5FF CRC64;

Query Match 14.7%; Score 494; DB 1; Length 558;
 Best Local Similarity 29.5%; Pred. No. 1,096-99;
 Matches 132; Conservative 89; Mismatches 182; Indels 45; Gaps 38;

DB 15 C-GPPR-DLPTALPASEMNOQDDESHHTLRNCRPGYRASSSSOSLYCKPQKW-QINIA 71
 OY 3 CKGPPRKNSEILSGSWSEQL-YSECTQATYKCRPGY-RTIGTIYKVKCN-GEWVSPNS 59
 DB 72 --CVKSCRNFGDLDNG--KYEVKTDFFGSOIEFSCGSEGYLLGSSSYSGEIOGKYSW 127
 OY 60 RICRRAPCGHGDIPFGSFRILAVGSEFFGAKVYITCDGTYLGE-IDYECDDAG--W 116
 DB 128 SDLPCEYACGMP-PDISNGK-HNGR-E-EFFETYSYTYKCDPFTLLGNASITCT 183
 OY 117 TNDIPICVWVC-LPVLELNGRIYSGAEPDOEYFGQVVRFCNSGFKIEGKEMHS 175
 DB 184 VVNTVGVWSPSPPCERITLCPMPKVLHGITNSGFKHYTKKDSVRYVCKOGFYLRGSGV 243
 OY 176 E-N---GLMSNEKQCVETISCLPPRVENGD-GIYKPYKNEFQYKCKOGFYKKERGD 230
 DB 244 IHCDAOSMSPVPCVCEINSCDIPDIPNALITSPRREKEDVYVGVTLRYCRPGYEP 303
 OY 231 AVCTGSG-WNPQSCHEMTCL-TYIPNG--IYTPH-RIKH--RIDDEIRIECKNGFYPA 283
 DB 304 TROMTYICQDLSMSLRGKEICCVPPDKSVYI-QHEK-AH-D-DNDCTY-FPGDE 358
 OY 284 TRSVVS-KCTIT-GWIPARCSLPCDFQPFHGRGLYEESSRPFPVPDYGKESYCDN 341
 DB 359 VSYICNDIMLT-ATCKSDGWHMPTSPCHOSODEPFAIHAHGRTRKSSSYVRYQV-YE 416
 OY 342 -GFTTSPQSYMDIYCTVNG-WEPEVP-CLROCIPIHYVEGESSYWMQR-YIEGOSAKV 397
 DB 417 CEESYRLV-GEATISCMYSCWTPAAPQC 443
 OY 398 CHSGYSLPNGODTYCTENGWSP-PRC 424

RESULT 6
 ID C4BP_BOVIN STANDARD; PRT; 610 AA.
 AC Q28065;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE C4B-BINDING PROTEIN ALPHA CHAIN PRECURSOR.

GN CABPA.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;
 OC Bovinae; Bos.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE-LIVER.
 RX MEDLINE: 95015909.
 RA Hillarp A., Thern A., Dahlbaeck B.;
 RT "Bovine C4b binding protein. Molecular cloning of the alpha- and
 beta-chains provides structural background for lack of complex
 formation with protein S.";
 RL J. Immunol. 153:4190-4199(1994).
 CC -1- FUNCTION: C4BP CONTROLS THE CLASSICAL PATHWAY OF COMPLEMENT
 ACTIVATION. IT BINDS AS A COFACTOR TO C3b/C4b INACTIVATOR
 (C3bINA), WHICH THEN HYDROLYZES THE COMPLEMENT FRAGMENT C4b. IT
 ALSO ACCELERATES THE DEGRADATION OF THE C4bC2a COMPLEX (C3
 CONVERTASE) BY DISSOCIATING THE COMPLEMENT FRAGMENT C2a. ALPHA
 CHAIN BINDS C4b. IT INTERACTS ALSO WITH SERUM AMYLOID P COMPONENT.
 CC -1- SUBUNIT: DISULFIDE-LINKED COMPLEX OF C4BP ALPHA AND BETA CHAINS
 (BY SIMILARITY).
 CC -1- SIMILARITY: TO C4BP BETA CHAIN AND TO PIG APOLIPOPROTEIN R.
 CC -1- SIMILARITY: CONTAINS 8 SUSHI (SCR) REPEATS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z31693; CAA83498.1; -
 DR HSSP: P10998; IYVC.
 DR PFM: PFO0084; SUSHI: 8.
 RW Complement pathway; Plasma; Glycoprotein; Repeat; Sushi; Signal.
 FT SIGNAL 1 48 BY SIMILARITY.
 FT CHAIN 49 610 C4b-BINDING PROTEIN ALPHA CHAIN.
 FT DOMAIN 49 542 8 X SUSHI (SCR) REPEATS.
 FT REPEAT 49 108 SUSHI 1.
 FT REPEAT 111 170 SUSHI 2.
 FT REPEAT 173 235 SUSHI 3.
 FT REPEAT 238 295 SUSHI 4.
 FT REPEAT 298 363 SUSHI 5.
 FT REPEAT 366 426 SUSHI 6.
 FT REPEAT 428 484 SUSHI 7.
 FT REPEAT 486 542 SUSHI 8.
 FT DISULFID 50 95 BY SIMILARITY.
 FT DISULFID 80 107 BY SIMILARITY.
 FT DISULFID 112 153 BY SIMILARITY.
 FT DISULFID 139 169 BY SIMILARITY.
 FT DISULFID 174 217 BY SIMILARITY.
 FT DISULFID 203 234 BY SIMILARITY.
 FT DISULFID 239 281 BY SIMILARITY.
 FT DISULFID 267 294 BY SIMILARITY.
 FT DISULFID 299 350 BY SIMILARITY.
 FT DISULFID 334 362 BY SIMILARITY.
 FT DISULFID 336 390 BY SIMILARITY.
 FT DISULFID 367 412 BY SIMILARITY.
 FT DISULFID 7402 425 BY SIMILARITY.
 FT DISULFID 429 471 BY SIMILARITY.
 FT DISULFID 457 483 BY SIMILARITY.
 FT DISULFID 487 528 BY SIMILARITY.
 FT DISULFID 514 541 BY SIMILARITY.
 FT DISULFID 549 549 INTERCHAIN (WITH BETA CHAIN) (POTENTIAL).
 FT DISULFID 561 561 INTERCHAIN (WITH BETA CHAIN) (POTENTIAL).
 FT CARBOHYD 66 66 POTENTIAL.
 FT CARBOHYD 221 221 POTENTIAL.
 FT CARBOHYD 525 525 POTENTIAL.
 FT CARBOHYD 602 602 POTENTIAL.
 SQ SEQUENCE 610 AA: 68886 MW: D806B270E8A06B58 CRC64;

Query Match 14.18, Score 474, DB 1; Length 610;
 Best Local Similarity 27.9%; Pred. No. 3,15e-94;
 Matches 120; Conservative 85; Mismatches 180; Indels 45; Gaps 35;
 Db 70 FENGTILRYCRPGYRIRSSRNFLICDGTDNW--KYEE-FCVKKRCNPGEL-LNG-OVI 124
 QY 24 YSEGTQATYKRCPGYRIRLUG-TIVVKCKNGE-WPDSNPRICRKRPGCHGDDTPPGSRFLA 81
 Db 125 VKTDSFGSEIEFSCSEGYVLIGSANSYCOLQDKGVWSDPLPQCIIAKCEPPTISNGR 184
 QY 82 VGSFEFGAKYVYTCDSGYQLGLGID-YRECDADG--WTNDIPICEVYKCLPVTLENGR 138
 Db 185 -HNGGDE-DF-YTYGSSVYVSCDPDFSMLGKASISCRVENKTTIVWSPSPSCKKVYCVQ 241
 QY 139 IVSGAAEPDQYHYGOVAVREFCNSGFKEGCKENKCS-EN--GLWSNKPQCVIEISCLP 194
 Db 242 PVMDKRTSGFPIITYQOSIVYACNKGFLBEGDSLHCADSNMNPPTCEINGCLG 301
 QY 195 PRVENGDIY-LKPYIKENEFQYKCKGKGYKRGDAVC-TGSGMNPQ-SCSEMCLT 251
 Db 302 LPHIPALMERYDHQYTDQVYDIGVLSYKCHFGYKRPETDGPITYVQSNLEMSPIYE 361
 QY 252 -PYIPNGI---YTPH-RIKHRI-DDE-I-RYCKNGFYPATRSPVS-KC-TITGWIPAPR 301
 Db 362 CKVCCPEPNLNTNGSITLH--RRP-----STSTCTYISGKISTECHKYMPDAL-CTK 414
 QY 302 CSLKCDPQPKRH-GRLYEESRRPFPVPGKEVSYCDNGFTTPSOS-Y-WDYLKCTV 358
 Db 415 HGTWSPRTPECRPDCKSPVIAHGKVVSKFTFDHQAVECDKGYLLGAKELT-CTS 473
 QY 359 NG-WEPEVP-CLRQCIFH-YVEYGESSYWRKRIEGQSAKQCHGYSILPNGODIYYCTE 415
 Db 474 SGWSPAVPQC 483
 QY 416 NGWSP-PRC 424
 RESULT 7
 ID CRI-HUMAN STANDARD; PRT: 2039 AA.
 AC P17927;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE COMPLEMENT RECEPTOR TYPE 1 PRECURSOR (C3b/C4b RECEPTOR) (CD35
 DE ANTIGEN).
 DE CRI OR C3bR.
 OS Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 89035992.
 RA Klinkstein L.B., Bartow T.J., Miletic V., Rabson L.D., Smith J.A.,
 RA Fearon D.T.;
 RT "Identification of distinct C3b and C4b recognition sites in the
 human C3b/C4b receptor (CRI, CD35) by deletion mutagenesis.";
 RL J. Exp. Med. 168:1699-1717(1988).
 RN [2]
 RP SEQUENCE OF 503-2039 FROM N.A.
 RX MEDLINE: 87168191.
 RA Klinkstein L.B., Wong W.W., Smith J.A., Weis J.H., Wilson J.G.,
 RA Fearon D.T.;
 RT "Human C3b/C4b receptor (CRI). Demonstration of long homologous
 repeating domains that are composed of the short consensus repeats
 RT characteristics of C3/C4 binding proteins.";
 RL J. Exp. Med. 165:1095-1112(1987).
 RN [3]
 RP SEQUENCE OF 761-783; 831-845 AND 1179-1195 FROM N.A.
 RX MEDLINE: 86067975.
 RA Wong W.W., Klinkstein L.B., Smith J.A., Weis J.H., Fearon D.T.;
 RT "Identification of a partial cDNA clone for the human receptor for
 RT complement fragments C3b/C4b.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:7711-7715(1985).

RT of the serum complement system."

RL Biochemistry 26:4668-4674(1987).

CC -1- ACTIVATION: C4BP CONTROLS THE CLASSICAL PATHWAY OF COMPLEMENT

CC (C3BINA), WHICH THEN HYDROLYSES THE COMPLEMENT FRAGMENT C4B. IT

CC ALSO ACCELERATES THE DEGRADATION OF THE C4Bc2a COMPLEX (C3

CC CONVERTASE) BY DISSOCIATING THE COMPLEMENT FRAGMENT C2a. ALPHA

CC CHAIN BINDS C4B. IT INTERACTS ALSO WITH SERUM AMYLOID P COMPONENT.

CC -1- SUBUNIT: HOMOEPTAMER; NOT COVALENTLY LINKED. MOUSE LACKS THE

CC BETA CHAIN OF C4BP.

CC -1- SIMILARITY: TO C4BP BETA CHAIN AND TO PIG APOLIPOPROTEIN R.

CC -1- SIMILARITY: CONTAINS 6 SUSHI (SCR) REPEATS.

CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-44 IS THE INITIATOR.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M17122; AAA37312.1; ALT_INIT.

DR PIR: A27117; NEMSC4.

DR HSSP: P10998; IYVC.

DR MGD: MGI:88329; C4BP.

DR PFW: PF00084; Sush1, 6.

KW Complement pathway; Plasma; Glycoprotein; Repeat; Sush1; Signal.

FT SIGNAL 1 56

FT CHAIN 57 469 C4B-BINDING PROTEIN.

FT DOMAIN 57 414 6 X SUSHI (SCR) REPEATS.

FT REPEAT 57 116 SUSHI 1.

FT REPEAT 119 177 SUSHI 2.

FT REPEAT 180 241 SUSHI 3.

FT REPEAT 244 300 SUSHI 4.

FT REPEAT 302 356 SUSHI 5.

FT REPEAT 358 414 SUSHI 6.

FT DISULFID 58 103 BY SIMILARITY.

FT DISULFID 88 115 BY SIMILARITY.

FT DISULFID 120 160 BY SIMILARITY.

FT DISULFID 146 176 BY SIMILARITY.

FT DISULFID 181 223 BY SIMILARITY.

FT DISULFID 209 240 BY SIMILARITY.

FT DISULFID 245 287 BY SIMILARITY.

FT DISULFID 273 299 BY SIMILARITY.

FT DISULFID 303 343 BY SIMILARITY.

FT DISULFID 329 355 BY SIMILARITY.

FT DISULFID 359 400 BY SIMILARITY.

FT DISULFID 386 413 BY SIMILARITY.

FT CARBOHYD 74 74 POTENTIAL.

FT CARBOHYD 227 227 POTENTIAL.

FT CARBOHYD 275 275 POTENTIAL.

FT CARBOHYD 292 292 POTENTIAL.

FT CARBOHYD 366 366 POTENTIAL.

FT CARBOHYD 381 381 POTENTIAL.

FT CARBOHYD 428 428 POTENTIAL.

SO SEQUENCE 469 AA; 51551 MW; 418137CB8D8C321 CRC64;

Query Match 12.9%; Score 434; DB 1; Length 469.

Best Local Similarity 31.0%; Pred. No. 2,23e-83;

Matches 116; Conservative 83; Mismatches 139; Indels 41; Gaps 36;

Db 58 C-GPPP-ALPNALPASDVRITDESHTTLKYCLDPGYGSGIRMMVYCCPSEGEWLSV-S 114

QY 3 CKGPPREMSILLSSMSQQL-YSEGTOATYKCRPGY-RTLTITVYKVC-NSEWPSNPS 59

Db 115 -CAKHCRCNPYLNG-Y-VN-GETTFGSGIIEFSQCEGFLVSSSTSSCEVRGKYAM 169

QY 60 RICRRRPGCHPDPTFGSGRIAVGSEFEAGARVYTCDEGYQLDEIDYR-ECADAG--W 116

Db 170 SNPEFACVYKGGPPDSDNGK-HSGT-E-DF-YYNNGICISTCPGRLVSGSPITGCTV 225

QY 117 TNDIPICVAVCLPTELENGRIYSGAAAPDOEYFGVAVRECSNGFKIEGKEMHCSE 176

Db 226 VKTVPWSSSPPECKICISQPNILHGVISGYKATYTHRDSVRLACNGTVLGRHVI 285

QY 177 -N-GL--WSNKKPCVCEISCLPVPVENGDGIY-LKPYKENEROYKOGFYKKEGDA 231

Db 286 ECGNGWMSLPTCE-EDCDLPPALVNGYYS-WYISKIT-LVYEDDKGRIYKRAIIS 342

QY 232 VCTSGSG-WNPQPSCEEMTC-LTPITPAGIYTPRIKRIHIDEIRYECKNFEYPATRSVS 289

Db 343 -CSFSKMKGTAPQCKA-LCKQPEVNGTLLS-DE-KDQY--VE-SENATIQDGSFALMG- 394

QY 290 KCTITGWI-PAPRCSLKPCCPPCKHRLYRESRRRYFPVPICKEYSYICDNGFTTSPQ 348

Db 395 S--OSISCSGCTVYPPVPRC 413

QY 349 SYWYILACTVNG-WEPEVP-C 367

RESULT 10

ID CFEH_HUMAN STANDARD; PRT; 270 AA.

AC P36980;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 01-JUN-1994 (Rel. 29, Last annotation update)

DE COMPLEMENT FACTOR H-LIKE PROTEIN DDESK59 PRECURSOR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Homiidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=LIVER;

RX MEDLINE; 92251200.

RA Skerka C., Timman C., Horstmann R.D., Zipfel P.F.;

RT Two additional human serum proteins structurally related to

RT complement factor H. Evidence for a family of factor H-related

RT genes.

RL J. Immunol. 148:3313-3318(1992).

CC -1- FUNCTION: MIGHT BE INVOLVED IN COMPLEMENT REGULATION.

CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.

CC -1- SIMILARITY: CONTAINS 5 SUSHI (SCR) REPEATS.

CC -1- SIMILARITY: STRONG, TO FACTOR H.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: X64877; CAA46096.1; -

DR PIR: S24564; S24564.

DR HSSP: P08603; IHP1.

DR PFW: PF00084; Sush1, 4.

KW Repeat; Glycoprotein; Sush1; Signal.

FT SIGNAL 1 18

FT CHAIN 19 270 POTENTIAL.

FT DOMAIN 19 270 COMPLEMENT FACTOR H-LIKE PROTEIN DDESK59.

FT REPEAT 19 85 SUSHI 1.

FT REPEAT 86 127 SUSHI 2.

FT REPEAT 128 205 SUSHI 3.

FT REPEAT 206 270 SUSHI 4.

FT CARBOHYD 126 126 POTENTIAL.

SO SEQUENCE 270 AA; 30651 MW; E1B2E9F139B217A8 CRC64;

Query Match 12.8%; Score 431; DB 1; Length 270;

Best Local Similarity 42.4%; Pred. No. 1.44e-82;

Matches 50; Conservative 22; Mismatches 46; Indels 0; Gaps 0;

Db 23 CDFPKNHGIYDEKYEFSQVPTGEVYFYSCYFNFYSPKSWTRITCAEESGMSPTPK 82

QY 307 CDFPKNHGIYDEKYEFSQVPTGEVYFYSCYFNFYSPKSWTRITCAEESGMSPTPK 82

Db 277 HSAKAFHOSCSFSCSEGFALVGP-EVVOCTASGWTAPAPACKAVOCCHLEAPSEG-T 334
 QY 81 AVGSFEFGKAVVYTCDEGQLEIDRECDAG-WNDIPICEVVKCLPVTLENGRI 139
 Db 335 MD-CVHPLTAFAVGSSECFECOPGVRGLDMLRCIDSGHMSADLPTEALSCPLESPV 393
 QY 140 VSGAAEPQOEYFVGVVAFECNSGFKIEGQKEMHCSEGLMSNKPQCVELSC-LPPRV 197
 Db 394 HGSMDCSPLRA-FOYDNTCSFRCAGFMRLG-ADIVACDNLGOWTAPAPQALQCODL 451
 QY 198 ENG-D-GIYLKPYKENERFOYKCKGQFVYKERGDV-CTGSG-WN-POPSCEMTCLTP 252
 Db 452 PVPNARNCNHPGAFVYGVSCFTCEGGLLVG-ASVLOCLATGNNSVPECOAIPC 510
 QY 253 YIPN-G-YTPHRI-KHRIIDEIRCKNGFPATRSVSKCTITG-WIPAPR-CSLKPC 307
 Db 511 T-PLSPONGT-MTCVOPLGSSSYKSTCOFTICDEGYSLSGPE-R-IDCTRSGRWTSPP 565
 QY 308 DEPOKRGRLYYEESRRFPYPIGKEYSYCDNGFTTPSGSYMDYLCTYNG-WEPEVP 366
 Db 566 MCEAIKCPFLPAPEEGSLDCSDTRGEFNVGSCNNGFRL-BGRNNVCTISGRMSA 624
 QY 367 -C-LRQCIFHYV-EVGESSYQWR--YLEGOSAKVQCHSGSYLPGODTYCTENG-WSP 420
 Db 625 TPPTC 629
 QY 421 -PPKC 424
 RESULT 13
 ID APOH MOUSE STANDARD; PRT: 345 AA.
 AC 001339:
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE BETA-2-GLYCOPROTEIN I PRECURSOR (APOLIPOPROTEIN H) (APO-H) (ACTIVATED
 DE PROTEIN C-BINDING PROTEIN) (APC INHIBITOR) (B2GPI).
 GN APOH OR B2GPI.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 92372000.
 RA Nonaka M., Matsuda Y., Shirosaki T., Moriwaki K., Natsuume-Sakai S.;
 RT "Molecular cloning of mouse beta 2-glycoprotein I and mapping of the
 RT gene to chromosome 11.";
 RL Genomics 13:1082-1087(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN-CBA/J; TISSUE-LIVER;
 RX MEDLINE: 94242017.
 RA Seiler G.C., Steel D.M., Zafirooulos A., Seery L.T.,
 RA Whitehead A.S.;
 RT "Characterization, expression and evolution of mouse beta 2-
 RT glycoprotein I (apolipoprotein H).";
 RL Biochem. Biophys. Res. Commun. 200:1521-1528(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN-BALB/C; TISSUE-LIVER;
 RA Kristensen T.;
 RT "Structure of the human beta-2-glycoprotein I gene.";
 RT Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: BINDS TO VARIOUS KINDS OF NEGATIVELY CHARGED SUBSTANCES
 CC SUCH AS HEPARIN, PHOSPHOLIPIDS, AND DEXTRAN SULFATE. MAY PREVENT
 CC ACTIVATION OF THE INTRINSIC BLOOD COAGULATION CASCADE BY BINDING
 CC TO PHOSPHOLIPIDS ON THE SURFACE OF DAMAGED CELLS.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- SIMILARITY: CONTAINS 4 SUSHI (SCR) REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).
 CC -----
 CC EMBL: D10056; BAA00945.1;
 DR EMBL: S70439; AAB30789.1;
 DR EMBL: Y11356; CAA72190.1;
 DR PIR: A43286; NBSM.
 DR HSSP: P10998; 1YVC.
 DR MGD: MGI:88058; APOH.
 DR PFW: PFW0084; sushi; 4.
 KW Heparin-binding; Glycoprotein; Plasma; Repeat; Sushi; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 345
 FT DOMAIN 22 261
 FT REPEAT 22 261
 FT REPEAT 83 138
 FT REPEAT 141 201
 FT REPEAT 204 261
 FT DOMAIN 263 345
 FT DISULFID 23 66
 FT DISULFID 51 79
 FT DISULFID 84 124
 FT DISULFID 110 137
 FT DISULFID 142 188
 FT DISULFID 174 200
 FT DISULFID 205 248
 FT DISULFID 234 260
 FT DISULFID 264 315
 FT DISULFID 300 325
 FT DISULFID 307 345
 FT CARBOHYD 162 162
 FT CARBOHYD 183 183
 FT CARBOHYD 193 193
 FT CONFLICT 252 252
 SQ SEQUENCE 345 AA; 38619 MW; C83F8A6EBD51C940 CRC64;
 Query Match 11.1%; Score 375; DB 1; Length 345;
 Best Local Similarity 29.3%; Pred. No. 1.37e-67;
 Matches 73; Conservative 46; Mismatches 115; Indels 15; Gaps 14;
 Db 41 YDPGQIYVSKPQGVSRGKRRFTCLTGMW-PINLR-CVPRVCFAGILENGIYRYT 98
 QY 24 YSEGTQATYKCRPGYKRLGLTVK-VCK-NEGWPSNPSRICKRPGCHPEDTPGSRFLA 81
 Db 99 -S-FEYPKNTSFACNPGF-PLNGTSSSKTEGKMSPDIPACARITCPPPYKFKALK 154
 QY 82 VGSEFEFGAKVYTCDESYQLGIDRECDAG-WNDIPICEVVKCLPVTLENGRI 140
 Db 155 DYRPSAGNNSLYDVTYFKCLPHFAMIGNDTWCETEGQNWTRL-PECLEYKCPPPPPEN 213
 QY 141 SGAAEPDQYFVGVVAFECNSGFKIEGQKEMHCSEGLMSNKPQCVELSC-LPPRVEN 199
 Db 214 GYVNPAPKPVLLYDKDAFFGCHETTKLDGPEAEACTGTGTSFLPTCRE-SCKLPRVKAT 272
 QY 200 GGCIY-LKPYKENERFOYKCKGQFVYKERGDVCTGSG-WNPOPSCEMTCLTPILNG 257
 Db 273 VLYGMRVK 281
 QY 258 I-YTPHRIK 265
 RESULT 14
 ID LEM3 RAT STANDARD; PRT: 768 AA.
 AC P98106:
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE P-SELECTIN PRECURSOR (GRANULE MEMBRANE PROTEIN 140) (GMP-140) (PADGEM)
 DE (CD62P) (LEUKOCYTE-ENDOTHELIAL CELL ADHESION MOLECULE 3) (LECAM3).
 GN SELP.

FT DISULFID 264 315
 FT DISULFID 300 325
 FT DISULFID 307 345
 FT CARBOHYD 149 149
 FT CARBOHYD 162 162
 FT CARBOHYD 183 183
 FT CARBOHYD 193 193
 FT CARBOHYD 253 253
 FT CARBOHYD 107 107
 FT VARIANT 266 266
 FT VARIANT 325 325
 FT VARIANT 335 335
 FT CONFLICT 121 121
 FT CONFLICT 188 188
 FT SEQUENCE 345 AA; 38298 MM; 63101704F8EDEF3F CRC64;

S -> N (IN APOH*1).
 /FTID=VAR_008169.
 V -> L (IN 23% OF THE POPULATION).
 /FTID=VAR_000673.
 C -> G (LOSS OF PHOSPHATIDYL SERINE-BINDING).
 /FTID=VAR_008170.
 W -> S (IN APOH*3W; LOSS OF PHOSPHATIDYL SERINE-BINDING).
 /FTID=VAR_008171.
 S -> C (IN REF. 7).
 C -> N (IN REF. 7).

Query Match 10.9%; Score 367; DB 1; Length 345;
 Best Local Similarity 28.5%; Pred. No. 1.79e-65;
 Matches 72; Conservative 54; Mismatches 106; Indels 21; Gaps 19;

Db 40 FYERGEIITSCRGYVSRGMRKFCPLTGLW-PINTLK-CTPRVCPFAGILENGAVRY 97
 QY 23 LISEGTOATKCRPGYRTLTGTVK-VCK-NGEWPSPNSRICKRRKPCGHPGDTPEFGSFR 80

Db 98 T--T-FEYPTNTISFSCNTGFIYNGA-DSAKTEBCKWSPELPVCAPICPPPSIPTATL 153
 QY 81 AVGSEFEFGAKVYVYTODEGYQLGEIDYRECDADG-WTNDIPICEVYKCLP--VTELENG 137

Db 154 RYKRPSSAG-NNSLYRDTAV-FECLPOHAMFGNDITICTHGMNTRL-PECREVKCPPSR 210
 QY 138 RIVSGAAEPDQVEYFGGVVREFCNSGFKIEGKEMHCSNGLWSNEKPKQCVETISC-LPR 196

Db 211 PDNGFVNYPAKFTLY-YKDKATFGCHDGYSLDPEIECTKLGNSMSPSCA-SCKVPV 268
 QY 197 VENGDDIY-LKP-VYKENERFQYCKOGFYKRGDAVCTGSG-WNPQPSCEWTCITLTPY 253

Db 269 KKATVYQGERVK 281
 QY 254 I-PNGIYTPHRIK 265

Search completed: Thu Jun 8 21:50:49 2000
 Job time : 19 secs.

 M E R E F

 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jun 8 21:51:06 2000; Maspar time 36.70 Seconds
 808.665 Million cell updates/sec

Tabular output not generated.

Title: >US-09-316-163-14
 Description: (1-428) from US09316163.pep
 Perfect Score: 3371
 Sequence: 1 EDCGKPPRENSEILSGSMS.....DTTYCTENGMSPPKCVRIK 428

Scoring table:
 PAM 150
 Gap 11

Searched: 225878 seqs, 69334122 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database:

sptremblj2
 1:sp-archae 2:sp-bacteria 3:sp-fungi 4:sp-human
 5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle
 9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified
 13:sp-vertebrate 14:sp-virus

Statistics: Mean 45.569; Variance 69.358; scale 0.657

Prod. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Prod. No.
1	2395	71.0	449	4	014570	COMPLEMENT FACTOR H PR	0.00e+00
2	1720	51.0	669	6	028085	CCP MODULES 3-12, WITH	0.00e+00
3	1136	33.7	452	11	061407	COMPLEMENT FACTOR H-RE	3.25e-267
4	1129	33.5	303	11	061405	COMPLEMENT FACTOR H-RE	2.70e-265
5	790	23.4	808	11	061408	COMPLEMENT FACTOR H-RE	4.80e-173
6	683	20.3	1053	13	091375	COMPLEMENT REGULATORY	2.44e-144
7	520	15.4	533	11	008569	ACROSOMAL MATRIX COMPO	3.02e-101
8	480	14.2	522	6	028769	COMPLEMENT RECEPTOR (F	7.77e-91
9	471	14.0	559	11	062315	COMPLEMENT REGULATORY	1.66e-88
10	470	13.9	657	4	014006	COMPLEMENT H FACTOR (F	3.02e-88
11	470	13.9	679	11	099254	COMPLEMENT RECEPTOR TY	1.25e-84
12	456	13.5	1911	6	029528	COMPLEMENT RECEPTOR 1	1.44e-82
13	448	13.3	331	4	092496	COMPLEMENT FACTOR H-RE	2.60e-82
14	447	13.3	343	11	061406	COMPLEMENT FACTOR H-RE	7.94e-83
15	449	13.3	661	6	029531	COMPLEMENT RECEPTOR 1	4.70e-82
16	449	13.3	2014	6	029530	COMPLEMENT RECEPTOR 1	4.70e-82
17	446	13.2	2039	4	016745	COMPLEMENT RECEPTOR 1	4.70e-82
18	445	13.2	2489	4	016744	COMPLEMENT RECEPTOR 1	4.70e-82
19	427	12.7	522	6	028797	UNKNOWN PROTEIN (FRAGM	6.37e-77
20	426	12.6	243	4	014310	FER-2 PRECURSOR.	6.37e-77

21	426	12.6	560	5	022328	COSMID T07H6.	6.37e-77
22	423	12.5	645	14	09WRU2	COMPLEMENT BINDING PRO	3.73e-76
23	419	12.4	315	6	028770	COMPLEMENT RECEPTOR (F	3.93e-75
24	411	12.2	579	11	060736	ZONA PELLUCIDA 3 RECP	4.32e-73
25	410	12.2	754	6	028290	CELL ADHESION MOLECULE	7.77e-73
26	409	12.1	497	11	063612	S12 ANTIGEN PRECURSOR.	1.40e-72
27	404	12.0	550	14	040912	ORF 04.	2.62e-71
28	400	11.9	550	11	035520	S12 ANTIGEN (FRAGMENT)	2.72e-70
29	400	11.9	417	11	035520	ORF 4.	2.72e-70
30	398	11.8	529	14	P87616	41KBP FRAGMENT FROM LE	8.76e-70
31	389	11.5	974	5	P91658	FURROEM.	1.68e-67
32	363	10.8	263	14	089859	HOMOLOG OF VACCINIA VI	5.99e-61
33	363	10.8	263	14	007033	D12L PROTEIN.	5.99e-61
34	356	10.6	263	14	089061	D15L.	3.40e-59
35	356	10.6	263	14	089076	B19L.	3.40e-59
36	359	10.6	740	4	095508	DJ780M13.1.2 (SELECTIN	6.05e-59
37	355	10.5	377	6	062838	MEMBRANE COFACTOR PROT	6.05e-59
38	355	10.5	378	6	062837	MEMBRANE COFACTOR PROT	6.05e-59
39	352	10.4	740	4	095507	DJ780M13.1.1 (SELECTIN	6.05e-58
40	347	10.3	285	6	019121	MEMBRANE COFACTOR PROT	6.05e-57
41	344	10.2	285	6	019127	MEMBRANE COFACTOR PROT	3.36e-56
42	344	10.2	285	6	019126	MEMBRANE COFACTOR PROT	3.36e-56
43	345	10.2	349	4	015429	CD46.	1.90e-56
44	341	10.1	369	6	P79138	MEMBRANE COFACTOR PROT	1.88e-55
45	334	9.9	483	11	064735	COMPLEMENT RECEPTOR RE	1.03e-53

ALIGNMENTS

RESULT	ID	1	PRELIMINARY:	PRT:	449 AA.
AC	014570;	P78435;			
DT	01-NOV-1996	(TREMblrel. 01, Created)			
DT	01-NOV-1996	(TREMblrel. 01, Last sequence update)			
DT	01-NOV-1999	(TREMblrel. 12, Last annotation update)			
DE	COMPLEMENT FACTOR H PRECURSOR.				
GN	HF OR CFH.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
OC	Eutheria; Primates; Catarrhini; Homiidae; Homo.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RK	MEDLINE; 88134059.				
RA	RIPACHE J., DAY A.J., HARRIS T.J.R., SIM R.B.;				
RT	"The complete amino acid sequence of human complement factor H.";				
RL	Biochem. J. 249:593-602(1988).				
RN	[2]				
RP	SEQUENCE OF 226-449 FROM N.A.				
RK	MEDLINE; 86169701.				
RA	KRISTENSEN T., WEISEL R.A., TACK B.F.;				
RT	"Structural analysis of human complement protein H: homology with Cdb				
RL	binding protein, beta 2-glycoprotein I, and the Ba fragment of B2.";				
RN	J. Immunol. 136:3407-3411(1986).				
RP	[3]				
RA	SEQUENCE OF 1-19 FROM N.A.				
RT	VIR D.P., WILLIAMS S.A.;				
RL	Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	SEQUENCE OF 1-9 FROM N.A.				
RA	DOMINGUEZ O.;				
RT	Thesis (1993), Immunologia, Hospital Tria's I Pujol, Spain.				
RL	EMBL; X07323; CA330403.1;				
DR	EMBL; M12383; AA52013.1;				
DR	EMBL; U56979; AAB01987.1;				
DR	EMBL; Z29665; CA82763.1;				
DR	HSSP; P10998; IYVC.				
DR	PFAM; PF00084; susbi; 7.				
KM	Signal.				
FT	SIGNAL.	1	18	POTENTIAL.	
FT	CHAIN	19	449	POTENTIAL.	
SEQUENCE	449 AA;	51033 MW;	61231E1B	CRC32;	
Query Match		71.0%;	Score 2395;	DB 4;	Length 449;

Best Local Similarity	67.4%;	Pred. No.	0.00e+00;
Matches	288;	Conservative	64;
		Mismatches	75

Matches	288;	Conservative	64;	Mismatches	75;	Indels	0;	Gaps	0;
---------	------	--------------	-----	------------	-----	--------	----	------	----

Db	19	EDCNBLPPRRNELLITGSMSSQOTPEGQALIKCRPGYRISGNVIMCRKGEWALNPLR	78
Oy	1	EDCKPPPRENSSELLTSGMSBQLVSEGTQATYKCRPGYRITGITVYCKKNGEWVPSNMR	60
Db	79	KQKRPPCGHPDPTPGTITLTGNWFEYGVAAVYTCNGYOLLEINVRREDDITDGMNDI	138
Oy	61	ICRRKPCGHPDPTPGSPRLAVGSFEFGAAVYVTCDEGTYLLDIEDIRCEDDAGMNDI	120
Db	139	PICEVYVKCLPTAPENKGIYVSSAEPDEYHFEGAVRVCVNSGKIEGDEBHCSDDGFW	198
Oy	121	PICEVYVKCLPATELENGRIYVGAAPDEYTFGGVYFECGSCGKIGOKEMHCSENGLW	180
Db	199	SKERPKCEIYSCSPDVYNGSPFIQKITYKNEFQYCKMNGYYSRGAVCTESGMRP	258
Oy	181	SNKRQCEIYSCLPYVENGSDITLKPYYKNEFRQYCKQGFYFKRGGAVCTGSGMNP	240
Db	259	LPSCBEKSCNDPYPNPGDYSPLRIKHTGDEITYQCRNGEYFATRGMTACTSTGWIPAP	318
Oy	241	QPSCEEMTCLPYPYIPNGIYTPHRIKHRIDEBYRECKNGEYFATRSPVSCITITGWIPAP	300
Db	319	RCTLKPCCDYPYIKHGSUYHEMRDPEYPPVAVNGKIYTYCDEHEPTBGSSTWDIHICTQDS	378
Oy	301	RCSLAKPCDFEYFKGRILYTESRNPYFVPLGKSYTYCNGGTTTPSQSWMDLRCITVNG	360
Db	379	WSPAVPCLAKCFYELTENGYNONGRKFEYVGKSIDVACHDIALPKQOTVVTOMENGWSP	438
Oy	361	WEPYVPCLRQCIHFHYVEYGESSWQRRYITBSQSAKVCHSGISYLPNQGDIYTYCTENGWSP	420
Db	439	TPRCIRY 445	
Oy	421	PKCVRI 427	

ID	RESULT	2
AC	Q28085	PRELIMINARY; PRT; 669 AA.
DT	01-NOV-1996	(TREMBLrel, 01, Created)
DT	01-NOV-1996	(TREMBLrel, 01, Last sequence update)
DT	01-NOV-1999	(TREMBLrel, 12, Last annotation update)
DE	CCP MODULES 3-12, WITH PARTS OF CCP 2 AND 13 (FRAGMENT).	
OS	Bos taurus (Bovine).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;	
OC	Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; Bovidae;	
RN	Bos.	
RP	[1]	
RC	SEQUENCE FROM N.A.	
RX	TISSUE=LIVER.	
RA	MEDLINE; 96202005.	
RT	SOMES C.J., DAY A.J., SIM R.B.;	
RT	"Prediction from sequence comparisons of residues of factor H involved	
RL	in the interaction with complement component C3b."	
DR	Biochem J 315:523-531(1996).	
DR	EMBL; X96697; CAAG7257.1; .	
DR	HSSP; P10998; IYVC.	
DR	PFAM; PF00084; Sushi; 11.	
FT	NON_TER	1
FT	NON_TER	1
SO	SEQUENCE	669 AA; 75683 MW; FAFODI74 CRC32;

```

Db      123  VLNQAVLPATRYQONRQVYRQCAAGGEYQROBDYTKSGMTPTCTCIEITDDPRIP 18
      156  RVENDDGITLPRVYKKNRFOYCKQOGVYKERDDAVCTSGNMPQSCCEMTLTIRYP 255
Qy
Db      183  NGVAPBELSKYRGODKITTECKKGFPEIRGTDATCTRGDGVAPPCAMKPCSPVYKHG 242
      256  NGIYTPHRIKIRIDEIETECKNGFFPATRSPVSKCITITGMPKRSKLPCDFPQCKHG 315
Qy
Db      243  RLYX--SYRGCFPARVNOQFYISCDHHFVPPSQRSMDHLACTAGWSPEEPCILQCIENY 300
      316  RLYYESRRPFEPVPIGEKYSYCDNCFPTPSQSYWYDLCTYVGMSEPEVPCILQCIENY 375
Qy
Db      301  LENGHNRREKRYLQGEFVARVHCYEGSLONDONTMTCTESGMSPPERCIRVX 353
      376  VEYGSSTWQNRRIYEGOSAKQOCHSGISLNGODDTYCTCTENGMSPPKCVRIK 428
Qy

```

```

ID      3
AC      061407
DC      061407.
DT      01-NOV-1986 (TRENBLER. 01, Created)
DT      01-NOV-1986 (TRENBLER. 01, last sequence update)
DT      01-NOV-1999 (TRENBLER. 12, last annotation update)
DE      COMPLEMENT FACTOR H-RELATED PROTEIN.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC      Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN      [1]
RP      SEQUENCE FROM N.A.
RA      VIR D.P., MUÑOZ-CANOVES P., KOZONO H., MARTIN L.G., TACK B.F.,
RA      CHAPLIN D.D.
RL      J. Biol. Chem. 0:0-0(0).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      VIR D.P., MUÑOZ-CANOVES P., KOZONO H., MARTIN L.G., TACK B.F.,
RA      CHAPLIN D.D.
RL      Submitted (APR-1990) to the EMBL/GenBank/DBJ databases.
DR      EMBL; M29010; AAA37415.1; -.
DR      HSSP; P08603; 1HT1.
DR      PFM; PF00084; sushi; 7.
SQ      SEQUENCE 452 AA; 51602 MW; 45C0BF61 CRC32;

```

ID	RESULT	4	PRELIMINARY:	PRT:	303 AA.
AC	061405				
AD	014005				
ADT	01-NOV-1996	(TRENDA1.01, Created)			
ADT	01-NOV-1996	(TRENDA1.01, Last sequence update)			
ADT	01-NOV-1996	(TRENDA1.12, Last annotation update)			
DE	COMPLEMENT FACTOR H RELATED PROTEIN.				
OS	MUS MUSCULUS (MOUSE).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia				

RX MEDLINE; 97284752.
 RA FOSTER J.A., FRIDAY B.B., MAULIT M.T., BLOBEL C., WINFREY V.P.,
 RA OLSON G.E., KIM K.S., GERTON G.L.;
 RT "Am67, a secretory component of the guinea pig sperm acrosomal matrix,
 RT is related to mouse sperm protein sp56 and the complement component 4-
 RT binding proteins.";
 RL J. Biol. Chem. 272:12714-12722(1997).
 DR EMBL; U75654; AAC13888.1;
 DR HSSP; P10998; IYVC.
 DR PFAM; PF00084; sushi; 7.
 DR Signal.
 FT CHAIN 1 28 POTENTIAL.
 FT SEQUENCE 29 533 ACROSOMAL MATRIX COMPONENT AM67.
 SQ SEQUENCE 533 AA; 59772 MW; 60E5A526 CRC32;
 Query Match 15.4%; Score 520; DB 11; Length 533;
 Best Local Similarity 27.2%; Pred. No. 3.02e-101;
 Matches 123; Conservative 105; Mismatches 179; Indels 46; Gaps 37;
 Db 29 DC-GPPPLFPASPVIOGY-ETNRTGTALKYNGHGYRWYSSHVLDINGSWI-VN-V 84
 QY 2 DCKPPPPREN-SEILSGMSSEQLYSEGTQATYKCRPGYRTLTGTYKVC-NGEWPSPNS 59
 Db 85 -FCAKKGRNGELANG-KWEITDLFGSTIEFSCSKYSLIGST-SCCESQGTVD 140
 QY 60 RICKRRGCHGHDTPFGSFLAVSGEFEGAKVYTDDEGYQLLEIDYRCDADG--- 115
 Db 141 WSDPLPCVYIKCDSPDISNGK-HSGTDE-DL-YTGSGLTYVCDPYSILGNASISCL 197
 QY 116 WTNDPICEVYKCLPVELENGRIVSGAEPDQYFYFGQVYFECNGSFKLEGKEMHC- 174
 Db 198 VAKTVGWSSNPTCEKVICRPHIKRGLFSGFGYTYTKTLVTSCKKGYLLRGSSI 257
 QY 175 -SEN-GLWSENERQCVESICLPRVENDGDIY-LKPYKENERFOYKCKGKGYKRGD 230
 Db 258 IHCEANSKWYPSIPTCPNGCIDPEVYIYSWENVLNKLNOELFEELGSLKYCKGIR 317
 QY 231 AVG-TSGGNPQ-PSCEMHC--LT--PIYP-N-GIYT-PIRIKHRIIDDEIRCKNGFY 281
 Db 318 PTNEPPTVTCQENLKAISKGCERVCPPNMEKRII-NE-RDFTGVYVAYEYIF 375
 QY 282 PATRSPVS-KCTIT-GWIPAPRCSLKPCDPFQKHSGLYEESSRPFPYPI-G-KEYSY 337
 Db 376 YMCDEGYPIISADGRS--SQADGMNPKPACESANCLPDLINGLSYEKDIYETEN 433
 QY 338 Y-CDNGFTTSPQSYWDYLRCTVNG-WEPEVP-CLRQ-CIFHYEYEGSSYWMORIRYEGQS 393
 Db 434 VTIHCDGYEVGPONII-CSENRTWTPEIPKC 465
 QY 394 AKVOCHSGYSLPNGQDPTYYCTEN-GMSPP-PKC 424
 RESULT 8
 ID 028769 PRELIMINARY; PRT; 522 AA.
 AC 028769;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
 DE COMPLEMENT RECEPTOR (FRAGMENT).
 OS Papio hamadryas (Hamadryas baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecoinae;
 OC Papio.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BONE MARROW;
 RA BIRNICHAM D.J., LOGAR C.M., SHEN X.P., CHEN W.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U77977; AAA99004.1;
 DR HSSP; P08603; IHFI.
 DR PFAM; PF00084; sushi; 7.
 DR NON_TER 1
 FT SEQUENCE 522 AA; 56626 MW; 7862072C CRC32;

Query Match 14.2%; Score 480; DB 6; Length 522;
 Best Local Similarity 33.2%; Pred. No. 7.77e-91;
 Matches 107; Conservative 74; Mismatches 102; Indels 39; Gaps 33;
 Db 68 FPGVLYKCELDPGVHG-KPESITCLNKSIVTSAK-DK-CTRKSCRNPKD-PVNGM-VHV 122
 QY 24 YSEGTQATYKCRPGYRTLTGTYKVC-NGEWPSPNSICKRRKRGHGDTPFGSFLAY 82
 Db 123 IKDIQFSQINYSCKNGRIIGSSSATCIISGNTVIMDNERTICEIIPGCPPT-IANGD 181
 QY 83 GSEFEFGAKVYITDDEGYQLLEIDYR-ECDADG-WTNDPICEVYKCLPVELENGR 138
 Db 182 FIS-TS--RETPFGSVYTRCNLGSGRKKLFELVEPSIYCTSKDQYIGSGAPAC 237
 QY 139 IVSGAEPDQY-FGQVYFECN--SGFK-I-E--GQKEMHC-S-EN--GLWSENERQCV 187
 Db 238 IIPKCMPPVENVLSVNSLSLNEVVEFRQCPGFYKRGPHVOCALNKWPELPS 297
 QY 188 VEIS-CLPRVENDGDIY-LKPYKENERFOYKCKGKGYKRGDAVCTG-SGNPQ-PS 243
 Db 298 CSRY-COPPEILHGEHTPSHODFSPQGEVYFSCPEP-Y-DLRGAASLHCTPOGWNPEA 354
 QY 244 CEEMTCLTP-YIPNGIYTPRIKHRIIDDEIRYCKNGFYPATSPVS-KCTITG-WIP-A 299
 Db 355 PICVYKSCDDEFLQPLHGRVLF 376
 QY 300 PROSLKPCD-F-PQFKHGRLYY 319
 RESULT 9
 ID 063135 PRELIMINARY; PRT; 559 AA.
 AC 063135;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
 DE COMPLEMENT REGULATORY PROTEIN.
 GN CRRY.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY;
 RA MEDLINE; 96006570.
 RA QUIG R.J., LO C.F., ALEXANDER J.J., SNEED A.E., MOXLEY G. III;
 RT "Molecular characterization of rat Crry: widespread distribution of
 RT two alternative forms of Crry mRNA."
 RL Immunogenetics 42:362-367(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-LIVER;
 RA DOHI N., SAKURADA C., NONAKA M., OKADA N., OKADA H.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U36532; AAA91821.1;
 DR EMBL; D42115; AAA22548.1;
 DR HSSP; P10998; IYVC.
 DR PFAM; PF00084; sushi; 7.
 SQ SEQUENCE 559 AA; 61680 MW; C87EEC58 CRC32;
 Query Match 14.0%; Score 471; DB 11; Length 559;
 Best Local Similarity 26.9%; Pred. No. 1.66e-88;
 Matches 123; Conservative 114; Mismatches 170; Indels 50; Gaps 36;
 Db 37 OCPAPPLFPYAKINPT-DESNFPGTSLKYRCRGYIKRQPSIT-CEVNSVWT-S-POD 92
 QY 2 DCKGPPPPRENSEILSGMSSEQLYSEGTQATYKCRPGYRTLTGTYKVC-NGEWPSPNSR 60
 Db 93 VCIRKQCEPLDPONG-I-VHVNTIRFGSSITTYCNGRYLIGSSAMCIIISDQSVAMD 150
 QY 61 ICRKRPCGHDTPFGSFLAVSGEFEGAKVYTCDEGYQLLEIDYR-ECDADG-GWT 117
 Db 151 AEAIPCEISPEIPIPSLNGDFES--PNRE-DFHYGVVYVQCNTDARGKFLNLVGEPS 207

```

QY 118 NDITCEVYKCLPYTELENGRIVSGAAPDDYFQGVVRECSN---G---FKIEGKE 171
DB 118 ICHTSIDQVWVSGPPEOCIELNKPETHENAVIVKNSLFLRDMVEFRODGFMM 267
QY 172 MHC-S-EN-GLMSNKKQCVIEIS-CLPFRYENGIGITL-KPYKENERFOYKCKOGVIV 225
DB 268 KGDSSVYCRSLNRREPOLPSCFKVKSAGALGELPNHGVFVPMNL-Q-LGAKVTFVNTG 325
QY 226 KERDVAVCTG-SGWNPPQ-PSC-EEWTC--LTPYIPNG-IYMPHRIKHRIIDEIRECKNG 279
DB 326 -YOLKGNSSSHCVLDGVESINSSVPCVOYICKLPOMSGFQKLOMKKDY--YCDN 381
QY 280 FYPTSPVSKCTITG---WIPA-PRCSLKPCCDFPKHRLYYEESRRPFVPIKKE 334
DB 382 VALCEDEGYTLEGGSSQSC-OSDAS-WDPLPKVSYQYICKLPOMSGFQKLOMKKDY 439
QY 335 YSYCDNGFTTSPQSYNDYLACTYNGMEPEVP-CLRQ--CLFHYVEYG-ESSYWORR-YI 389
DB 440 YGDVVALECEDEGYTLEGGSSQSCOSDASWDPPLPKCV 476
QY 390 EGOSAKVQCHSGSYSLPNQDITYCTENGWSP-PCV 425

RESULT 10
ID Q14006 PRELIMINARY: PRT: 657 AA.
AC Q14006;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE COMPLEMENT H FACTOR (FRAGMENT).
GN HP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 88025472.
RA DAY A.J., RIPOCHE J., LYONS A., MCINTOSH B., HARRIS T.J., SIM R.B.;
RT "Sequence analysis of a cDNA clone encoding the C-terminal end of
RT human complement factor H."
RL Biosci. Rep. 7:201-207(1987).
DR EMBL: M17517; AAA52016.1; -.
DR HSSP: P08603; 1HFH.
DR PFAM: PF00084; sushi; 11.
FT NON_TER 1
SQ SEQUENCE 657 AA; 74247 MW; F4AB5238 CRC32;

Query Match 13.9%; Score 470; DB 4; Length 657;
Best Local Similarity 26.0%; Pred. No. 3.02e-88;
Matches 107; Conservative 78; Mismatches 201; Indels 26; Gaps 25;

DB 197 FDHNSNIRYROR-G-KE-GWITVCINGRMDPEVNCMAQIOLCPPPPQIP-NSHNMTTT 252
QY 24 YSEGTQATYCRPEYRGLTIVKCKNGEWPNSPSCRKRRPGHGIDTFFGFRILAV 83
DB 253 LNRVDEKSVYLCQENT-LIQEGEITCKDGRW-OSIPLCEKIKPCQOPDIEHGTINS 310
QY 84 SEFEFGAKVYVYTCDEGYQLGELIDYRECDADGWTNDIPIC-EYVCKLPVTELENGRIYSG 142
DB 311 RS-SQESYAGHTKLSYCEGFRISENETTCY-MGKMSSP-POCEGLPKCSPEISHGV 367
QY 143 AAEPDQYTYGGVYRFECSNGFKIEGKEMHCSENGLSMNSKPCVETISCL-PRVINGD 201
DB 368 VAHMSDSYQGEVYTKCFEGFIDGPAIAKCLGEMSHSPSCIKTDLSPSENAIPM 427
QY 202 GYLKPYKRENERQYKCKGFGYKRGDAVCTSGSNWPPDSCHEMCTLT-PYIPNGIYT 260
DB 428 GEKKDYKAAEQYTYT-ATYYKMGDSANTCINSMTGRTCDTSCVNDPTVQNA-YI 485
QY 261 PHRTK-HRIDETIEYCKNKGYPATRSVSKCTITGWIIPAPRCSLKPCCDFPKHGRILY 319
DB 486 VSRQMSKYP-S-GERVYQCRSPYEMFGDE--EYV-CLNGWMTPEPPQCKDSTGKCPFP 540

```

```

QY 320 EESRRPYVPVIGKKEYSYXCNDNGFTTSPQSYNDYLCTVNGW-EP-EV-PCLRQC-IFHY 375
DB 541 IDNGITSEPLSYAPASSVYCOONLYQL-EGNKRITCRNGOWSEPPKCH 591
QY 376 VEYGE-SSYWMORRYIEGOSARVQCHSGSYSLPNQDITYCTENGWSPPCV 426

RESULT 11
ID Q09254 PRELIMINARY: PRT: 679 AA.
AC Q09254;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE COMPLEMENT RECEPTOR TYPE 2 PRECURSOR (CR2)
DE (ALTERNATIVELY SPLICED VERSION) (FRAGMENT).
GN CR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-BALB/C.
RX MEDLINE: 90229754.
RA KURTZ C.B., O'TOOLE E., CHRISTENSEN S.M., WEIS J.H.;
RT "The murine complement receptor gene family. IV. Alternative splicing
RT of Cr2 gene transcripts predicts two distinct gene products that share
RT homologous domains with both human CR2 and CRL1."
RL J. Immunol. 144:3581-3591(1990).
RN [2]
RP SEQUENCE OF 21-367 FROM N.A.
RX MEDLINE: 95105691.
RA KIM Y.U., KIMOSHITA T., MOLINA H., HOURCADE D., SEYA T., WAGNER L.M.,
RA HOLERS V.M.;
RT "Mouse complement regulatory protein Crry/p65 uses the specific
RT mechanisms of both human decay-accelerating factor and membrane
RT cofactor protein."
RL J. Exp. Med. 181:151-159(1995).
CC -1- ALTERNATIVE PRODUCTS: ALTERNATIVE SPLICING OF CR2 GENE TRANSCRIPTS
CC PREDICTS TWO DISTINCT GENE PRODUCTS THAT SHARE HOMOLOGOUS DOMAINS
CC WITH BOTH HUMAN CR2 AND CRL1.
CC -1- TISSUE SPECIFICITY: B LYMPHOCTES.
DR EMBL: U17128; AAA78271.1; -.
DR EMBL: U17123; AAA78271.1; JOINED.
DR EMBL: U17124; AAA78271.1; JOINED.
DR EMBL: U17125; AAA78271.1; JOINED.
DR EMBL: U17126; AAA78271.1; JOINED.
DR EMBL: U17127; AAA78271.1; JOINED.
DR EMBL: M36470; AAA37449.1; -.
DR HSSP: P10998; 1VVC.
DR MGD: MGI:88489; Cr2.
DR PFAM: PF00084; sushi; 10.
DR KW Complement pathway; Glycoprotein; Transmembrane; Repeat; Signal;
KW Alternative splicing; Receptor.
FT SIGNAL 1
FT CHAIN 12 679 COMPLEMENT RECEPTOR TYPE 2.
FT NON_TER 1
SQ SEQUENCE 679 AA; 74916 MW; BF08AC75 CRC32;

Query Match 13.9%; Score 470; DB 11; Length 679;
Best Local Similarity 31.7%; Pred. No. 3.02e-88;
Matches 101; Conservative 72; Mismatches 111; Indels 35; Gaps 25;

DB 42 FAITWEYCRPGYFRKSFIT-CLFTSKW--SDAOPFKRRCMPPOEILHGSVHINT 98
QY 24 YSEGTQATYCRPEYRGLTIVKCK-NGEWPNSPSCRKRRPGHGIDTFFGFRILAV 82
DB 99 G-LEFGSTTYSNOGRRLIGDSATCIYSDNTVMNDMPPLCESIPCESPAISNGDF 156
QY 83 GSEFEFGAKVYVYTCDEGYQLGE--IDYRECDADG-WTNDIPICEVYKCLPYTELENGRI 139
DB 157 YS-SS-RD-SFFIGVYTYTCHGKNREKLFDLVGEESIYCTSKDNQVGIWNSDPPOCIP 213

```

QY 140 VSGAEPDQEFYFGVYREFCNSG-----FKIGOKEMHC-S-EN--GLMSNEKPCQVE 189

Db 214 RYKCPMEIENGVLSEGP-KHSEFLNDTVIFKCKSGFTMKSGRIAKQCPNKSPEPLPTC 272

QY 190 -ISCLPRVENG--DGILYKRVYKENERFOYCKCKGQFVYKBERGAVC-TGSGMPO-PSC 244

Db 273 F-MGCLPPOINTLHGDKKDEFFSVGOKVSYTCNPG-YTLIGTILVECTSLGTWNTVPT 330

QY 245 EEMTCLP-YIPNGIYPRRIKHRIIDELRECKNGFYPAIRSVSKCTTG-WIPA-PR 301

Db 331 CEVKSCDAPNHLHGRVF 349

QY 302 CSLKPCD-FP-QFKHGRLY 318

RESULT 12 PRELIMINARY; PRT; 1911 AA.

ID Q29528; AC Q29528; DT 01-NOV-1996 (TREMblrel, 01, Created) DT 01-NOV-1996 (TREMblrel, 01, Last sequence update) DT 01-NOV-1999 (TREMblrel, 12, Last annotation update) DE COMPLEMENT RECEPTOR 1 (FRAGMENT). GN CRI. OS Papio hamadryas (Hamadryas baboon). OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; OC Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; OC Papio. RN [1]. RP SEQUENCE FROM N.A. RA CLEMENZA L., SUBRAMANIAN B.V., NICKELLS M.W., HOURCADE D.E., RA ATKINSON J.P., RL Submit (MAR-1995) to the EMBL/GenBank/DBJ databases. DR EMBL: L39791; AAA62170.1; -. DR HSSP: P08603; IHCC. DR PIRAM: PF00084; sushl; 29. FT NONTER 1911 1911 SQ SEQUENCE 1911 AA; 210173 MW; 030640C2 CRC32;

Query Match 13.5%; Score 456; DB 6; Length 1911; Best Local Similarity 31.8%; Pred. No. 1,25e-84; Matches 118; Conservative 81; Mismatches 128; Indels 44; Gaps 37;

Db 850 FPIGTSLKCYGKREYTG-KPFSITCLD-NLYWSSPKDYCKKSKCTPPD-FVNGM-VHYI 905

QY 24 YSEGTQATYKRCPRGRLGTIVKCKNGEWPNPSNRKCRKPCGHPDTPFGSFRILAVG 83

Db 906 TDIVGSRINSCCTGHLGHSS-AECIIIGNTAWSTKPCRICPQLP-PIIANGD 963

QY 84 SEFEFGAVVYTCDEGYLLGEIDYREC--DAD-G-WINDIPICVYVC-LPYTELENGR 138

Db 964 FIS--T-NREYFHYGVVYRCNLGSGRKLFEVGEPSIYCTSKDQVGIWGPAPQC 1019

QY 139 IVSGAEPDQGY-FGVYVREC--SGFK-I-E--GQKEMHC-S-EN--GLMSNEKPCQ 187

Db 1020 IIPKCMPTNENGLVSVNSLSFLNEVEFRQCPGVYMGPRVQCALNKWPELPS 1079

QY 188 VEIS-CLPVRVNGDGYL-KPYKENERFOYCKCKGQFVYKBERGAVC-TGSGMPO-PSC 243

Db 1080 CSRV-COPPELHGHETPSHODKSPQCEVYSCPEP-Y-DLRGAALHCTPOQDMPE 1136

QY 244 CEEMTCLP-YIPNGIYPR-HRIKHRIIDELRECKNGFYPAIRSVSKCTTG-WIPA-PR 298

Db 1137 APCAVKSCD--DF-LGOLHGRVLP-FNLQAGKVSFVDEGFRILKSSVSHCVLGM 1192

QY 299 APRCSLKPCDPPQFKHGRLYEESRRPYFPVPIGKEYSYCDNGFTTSSQSYMDLACTV 358

Db 1193 RSLMNNNSVPC 1203

QY 359 NG-WEPEVP-C 367

RESULT 13 PRELIMINARY; PRT; 331 AA.

ID Q92496; AC Q92496; DT 01-FEB-1997 (TREMblrel, 02, Created) DT 01-FEB-1997 (TREMblrel, 02, Last sequence update) DT 01-NOV-1999 (TREMblrel, 12, Last annotation update) DE COMPLEMENT FACTOR H-RELATED PROTEIN 4 PRECURSOR. OS Homo sapiens (Human). OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; OC Eutheria; Primates; Catarrhini; Homnidae; Homo. RN [1]. RP SEQUENCE FROM N.A. RA SKERRA C., HELLMAGE J., WEBER W., TILKORN A., BUCK F., MARTI T., RA KAMPER E., BEISIEGEL U., ZIPFEL P.F.; RT "The human factor H-related protein 4 (FHR-4). A novel short consensus repeat-containing protein is associated with human triglyceride-rich lipoproteins." RT J. Biol. Chem. 272:5627-5634(1997). DR EMBL: X98337; CAA66980.1; -. DR HSSP: P10998; IYVC. DR PIRAM: PF00084; sushl; 5. KW Signal. FT SIGNAL. 1 19 POTENTIAL. FT CHAIN 20 331 COMPLEMENT FACTOR H-RELATED PROTEIN 4. SQ SEQUENCE 331 AA; 37325 MW; 2505D66C CRC32;

Query Match 13.3%; Score 448; DB 4; Length 331; Best Local Similarity 43.6%; Pred. No. 1,44e-82; Matches 61; Conservative 22; Mismatches 52; Indels 5; Gaps 4;

Db 9 LTLWVSCANGQEVKPCDPEIIOHGLYKSLRLYFPAAGSYCYCDONFVTPSGSYW 68

QY 293 ITGWIAPARC-SLKPCDFQFKHGRLYEESRRPYFPVPIGKEYSYCDNGFTTSSQSYW 351

Db 69 DYHCTODGMLTPYPCICRTSKSDIEENGFISSSIYLNKRIQYCKPGYATAAGNS 128

QY 352 DYLCITVNGMEPEVPCLRQCI-FHY-VEYGBSSYWRRIEGBQAKYCHSGISLPGQD 409

Db 129 SGTITCLONGMSAOPICIKF 148

QY 410 T-YCTENGWSPPKCVRI 427

RESULT 14 PRELIMINARY; PRT; 343 AA.

ID Q61406; AC Q61406; DT 01-NOV-1996 (TREMblrel, 01, Created) DT 01-NOV-1996 (TREMblrel, 01, Last sequence update) DT 01-NOV-1999 (TREMblrel, 12, Last annotation update) DE COMPLEMENT FACTOR H-RELATED PROTEIN. OS Mus musculus (Mouse). OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. RN [1]. RP SEQUENCE FROM N.A. RA MEDLINE: 90153969. RA VIK D.P., MONOZ-CANOVES P., KOZONO H., MARTIN L.G., TACK B.F., RA CHAPLIN D.D.; RT "Identification and sequence analysis of four complement factor H-related transcripts in mouse liver." RT J. Biol. Chem. 265:3193-3201(1990). DR EMBL: M29008; AAA37414.1; -. DR HSSP: P08603; IHFI. DR PIRAM: PF00084; sushl; 4. SQ SEQUENCE 343 AA; 38443 MW; 6CFD5090 CRC32;

Query Match 13.3%; Score 447; DB 11; Length 343; Best Local Similarity 41.4%; Pred. No. 2,60e-82; Matches 55; Conservative 29; Mismatches 49; Indels 0; Gaps 0;

Db 14 LTSWFSTAKGEVSLCDPFIKRIHGRILYDEKNEPSSVSGKILYSCYNFAPSPNSFWT 73

QY 293 ITGWIAPARC-SLKPCDFQFKHGRLYEESRRPYFPVPIGKEYSYCDNGFTTSSQSYW 352

Db 74 RITCESGSPYKCLRLCEFPVENGNSTSGQTHVEGDIYQVVCNOGYSLONNOSTIT 133
Y 353 YLRCTVNGWEPEVPCLRQCFHFHYVEGESSYQRRYIEGSAKVOCHSGXSLPNGDITY 412
Db 134 CAEGMSITPKCI 146
Y 413 CTENGMSPPKCV 425

RESULT 15
ID 029531 PRELIMINARY: PRT: 661 AA.
AC 029531:
DT 01-NOV-1996 (TIREMBLrel. 01, Created)
DT 01-NOV-1996 (TIREMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TIREMBLrel. 12, Last annotation update)
DE COMPLEMENT RECEPTOR 1 PRECURSOR (FRAGMENT).
GN CRL.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Pan.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94292799.
RA BIRMINGHAM D.J., SHEN X.P., HOURCADE D., NICKELS M.W., ATKINSON J.P.;
RT "Primary sequence of an alternatively spliced form of CRL Candidate
RT erythrocytes.";
RL J. Immunol. 153:691-700(1994).
DR EMBL: L24921; AAA51439.1; -.
DR HSSP: P10998; IYVC
DR PFAM: PF00084; sush1; 9.
KW Signal; Alternative splicing.
FT NON_TER 1
FT SIGNAL 1
FT CHAIN 17
FT NON_TER 17
SO SEQUENCE 661 AA; 72966 MM; 9D78E262 CRC32;

Query Match 13.3%; Score 449; DB 6; Length 661;
Best Local Similarity 32.3%; Pred. No. 7.94e-83;
Matches 104; Conservative 73; Mismatches 107; Indels 38; Gaps 33;

Db 38 EPIGTYLNECRPGYGR-PFSIICLNKSVWTGAK-DR-CRKSQRNPPD-PVNGM-VHV 92
Y 24 YSEGTQATYKCRPGYGRITLGTIVKVC-KNGEWYSPNSRICRRKRGHPGDTFFGSRFLAV 82
Db 93 IKDIOFGSQIKYSCITKGYRLIGSSSATCIISGDTYIMNEMPICDRIPOGLPPT-ITNGD 151
Y 83 GSEFFFGAKVYITCDEGYQLGEIDYR-ECDADG--WTNDIPICEVAKC-LPYTELENGR 138
Db 152 FIS--TNRE-NFHGVSVTYRCNPGSGGRKYFELVEPSIYCTSDNDQVGIWSPAPOCI 208
Y 139 IVSGAABPDQYFYFGVYFECN--SGF-KI-E--GQKEMHC-S-EN--GLMSNEKPCV 188
Db 209 IPNCTIPNVNENGLIVSNBSLFLSNEVEYERCPGPFVAKGPPRYKQALNKWEPDLPSC 268
Y 189 EIS-CLPRVENGDGILY-LKPYKENERFOYKCKOGFYKKEGDAVCTG-SGMNPO-PSC 244
Db 269 SRV-CQPPDVLHAERTORDNFSGOEVFYSCEPG-Y-DLRGAASLRCTPOGDWSPAT 325
Y 245 EEMTCLTP-YIPNGIYTPH-RIKHRIDDEIRYECNNGFYPATRSPVS-KCITTG-WIPA- 299
Db 326 PTCEYKSCDDFMGOLLNGRLVLE 347
Y 300 PRCSLKPDCD-F-POFKHGRLY 319

Search completed: Thu Jun 8 21:51:46 2000
Job time : 40 secs.

